Higgs Boson Decays to $b\bar{b}$ with associated t-quarks with the ATLAS detector at the Large Hadron Collider

A thesis submitted to the University of Glasgow in fulfilment of the requirements for the award of the degree of Doctor of Philosophy in the College of Science and Engineering

Azzah Al-Shehri

School of Physics and Astronomy

University of Glasgow

April 2018

Abstract

This thesis uses 32.88 fb^{-1} of pp collision data gathered at the LHC by the ATLAS detector during 2016 at $\sqrt{s}$ =13 TeV. The analysis employs kinematic fitting techniques by applying the KLFitter package on the signal-rich region using only the 6 jets selection mode (kB6). It construction variables providing good separation between signal and background in the search for $t\bar{t}H(H \to b\bar{b})$ in the single-lepton final state (electron or muon). The scalar sum of transverse momenta is the variable of choice for the fit in signal-depleted regions. Using Boosted Decision Trees in the fit of signal-rich regions, a 95% CLs exclusion limit (significance) of 5.4 (4.25 σ) is obtained, with the corresponding ratio of the measured $t\bar{t}H$ signal cross-section to the Standard Model expectation of $3.69^{+0.98}_{-0.88}$. This result indicates that there is an excess of events above the background expectation for the SM Higgs boson with mass of 125 GeV. The excess is even greater than the SM would predict (μ =1). This excess over the SM prediction could be interpreted as a statistical fluctuation, and is not significant. More data would likely moderate this statistical aberration.

Declaration

I declare that, except where otherwise stated, this thesis presents work completed by me, Azzah Alshehri, between 2013 and 2018 in the Experimental Particle Physics research group in the School of Physics and Astronomy at the University of Glasgow. None of the work described here has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

The analysis of the output presented in Chapters 7, 8, and 9, is performed using custom code which is written from scratch and is my own, except where otherwise stated.

— **Azzah Alshehri, 2018-07-12**

Copyright

Acknowledgements

Thanks is but one word, and behind it there is a lot of gratitude to all those whom these words cannot thank enough for their support. Each of their names engraved in my heart and my memory like the Acknowledgements in bold font at the top of this page.

First of all I want to thank my supervisor, Tony Doyle, for all the hard work, support and guidance. I learned from you the seriousness of working together with teamwork and the simplicity of your meticulous handling of every physics issue. A big thank you also goes to Christopher Collins Tooth who gave me all his experience in physics with all generosity and kindness. He was a friend, guide and motivated in every time I was frustrated. I also want to thank warmly, Andrea Knue for her support and help to get this work to completion. Thank you Andrea so much! Thank you also to Sarah Boutle who accompanied me from the beginning of the journey to its end. Thanks a lot to Will Breaden Madden, Gavin Kriby, and all the ttH group and particle physics groups in Glasgow university.

Thanks with no limits to my parents who supported me from the first day of my life until I completed my PhD thesis. I present them all my achievements. I learned from you the love of knowledge and strive to reach our goals. Thank you in all meaning of love, appreciation and gratitude to my friend, lover, soul mate and life mate, my dear husband Saleh. Thanks Saleh for the extreme care that you gave me and our children during my exhausting years of research. Thanks to my children, Raya, Yazeed, Jalal and Maria for your patience since I was away and busy during the last five years. A big thank you also to my sisters and brothers, Sarah, Bathan, Azizah, Loulah, Nouf, Meshal and Mishael. I am lucky and happy because you are always around me with your support and love even when I am living abroad for a long time.

Thanks to all my friends here in the UK and in Saudi Arabia, I do not even have to mention your names but i will never forget your support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Higgs boson is an important piece of the Standard Model (SM) whose discovery was announced by ATLAS [1] and CMS [2] in July 2012 (forty years after its prediction) and is considered a milestone for high-energy physics. The discovery of the Higgs boson with a mass of ∼125 GeV was the primary success of explorations at the Large Hadron Collider (LHC). This discovery, in addition to the discovery of the vector bosons ($W^{\pm}$ and Z) [3, 4] with the expected properties, increased our confidence in the electroweak sector of the SM. The origin of masses of the particles is described via the Higgs mechanism [5], which was an ideal solution responsible for electroweak symmetry breaking. The mass of all the particles of the model are generated by the Higgs mechanism and at the same time this creates an associated particle, the Higgs boson. The measurement of all the properties of this new particle is essential for a confirmation of the Standard Model or for the discovery of new physics processes. Of particular interest are the Yukawa couplings, which in the Standard Model can be inferred from the measured fermion masses. The top Yukawa coupling (Y_t) to the fermions has a unique feature that its value is close to unity, which is the largest among the fermions in the SM and could point to the presence of new physics. So, the exact measurement of Y_t means that channels involving a Higgs boson which couples to top quarks are fundamentally interesting.

Access to Y_t is generated in the LHC through two main processes; gluon fusion involving a fermion loop, which complicates the extraction of the Yukawa couplings, and the $t\bar{t}H$ direct production channel which also allows measurement of top and b-quark Yukawa couplings to the Higgs field. The $t\bar{t}H$ has another unique feature: the increase of energy at LHC in Run2 increases the cross-section of this channel as a function of the centre-of-mass energy of the pp collisions faster than the other Higgs production channels. The SM Higgs production cross sections at $\sqrt{s}$ =7-13 TeV

increase by a factor five from Run 1 to Run 2 [6]. This leads to a better precision on the upper limit of the signal strength $\mu = \sigma^{\mathrm{ttH}}_{observed}/\sigma^{\mathrm{SM\ ttH}}_{expected}$. The ATLAS Run 1 (2012) search results for $t\bar{t}H(H \to b\bar{b})$ in the single lepton channel yielded a signal strength measurement of $\mu = 1.5 \pm 1.1$ with an observed (expected) limit of 3.4 (2.2) [7]. This was extended in Run 2 to include both single and dilepton channels and this gave a combined signal strength parameter measurement of $\mu = 0.84^{+0.64}_{-0.61}$, corresponding to an observed (expected) significance of $1.4\sigma(1.6\sigma)$ [8].

However, the $t\bar{t}H$ channel has a small cross-section due to the large invariant mass of the final state objects, such that it contributes only around 1% to the total Higgs boson production cross-section [9, 10].When we consider the decay $H \to b\bar{b}$, which is the largest branching ratio for the Higgs boson (58%), in association with $t\bar{t}$ pair, the channel becomes less difficult but still very challenging.

The $t\bar{t}H(H \to b\bar{b})$ channel has different signatures which are distinguished by the type of top quark decay mode. The top quark decay probability of $t \to Wb$ is almost 100% and then the final state of $t\bar{t}H(H \to bb)$ is divided into three classes according to the W boson decay.

The analysis presented in this work concerns the channel in which the Higgs boson decays to a pair of b quarks and the $t\bar{t}$ system decays semileptonically, specifically electron+jets and muon+jets in ATLAS at 13 TeV using the full 2016 LHC data set. Exclusive regions are defined according to the number of jets, and number of b-tagged jets at the 77% b-tagging efficiency working point.

The search for the $t\bar{t}H(H \to bb)$ process faces several difficulties because of the large backgrounds, low cross-section compared with $pp \to VH$ or $pp \to H$ channels and its final state is complicated by the presence of 6 high-p_T jets, two of which are light-jets and the remaining 4 jets are the b-quark jets. This gives rise to a challenging analysis in two ways. The high combinatorics from the 4 b-jets plus possible mis-tagging of light-jets, and the large backgrounds from $t\bar{t}$+light jets. The result is the small signal to background ratio after the event selection and categorisation as signal-rich and background-enriched regions. In order to increase the purity of the signal-rich regions; $(\geq 6j, \geq 4b)$, defined here as $(\geq 6j, 3b)$ and $(5j, \geq 4b)$, it is the aim of this analysis to optimize the sensitivity of the search using a kinematic fit (KLFitter) which tries to reconstruct the Higgs boson, leptonic and hadronic top quarks and the W bosons from the measured objects. In doing so, it provides some output variables that are directly sensitive to how much an event looks like the signal. For the final measurement these additional powerful discriminating variables can then be combined with other non-KLFitter variables in a MultiVariate Analysis

by providing four *Boosted Decision Trees* (BDTs) [11] according to the four signal regions which are used.The single KLFitter signal region ($\geq 6j, \geq 4b$), and the three "non-KLFitter" signal regions; ($\geq 6j, \geq 4b$), ($\geq 6j, 3b$) and ($5j, \geq 4b$). Then, these signal-rich regions are exploited with a likelihood fit to evaluate the signal strength μ. In doing the fit, a 95% confidence-level exclusion limit is obtained, associated with the signal strength. Many systematic uncertainties are considered from several sources: from the theory, from object reconstruction techniques and from the detector.

This thesis is organized as follows, Chapter 2 focuses on the theoretical basis of the Standard Model, and introduces the Higgs mechanism and Higgs boson physics. In Chapter 3, the ATLAS detector is described in detail. Chapter 4 describes the categorization of events, the MultiVariate Analysis and the Statistical tools that are used in the analysis in terms of profile likelihood fit. The reconstruction and selection of individual physics objects are summarized in Chapter 5. A brief description of the data that are used by this analysis and the signal and background modelling are presented in Chapter 6. In Chapter 7, the implementation of the kinematic fit to the $t\bar{t}H(H \to b\bar{b})$ semi-leptonic channel is detailed, followed by the validation of this implementation and the study of the performance of the fit in Chapter 8. In Chapter 9, the systematic uncertainties are discussed and the output variables of the KLFitter are combined with the non-KLFitter ones, in order to discriminate signal from background. The final results on the signal strength and the upper limits are then presented.

Chapter 2

The Standard Model

2.1 Introduction

This chapter introduces a brief summary of the Standard Model (*SM*). Then, the description of how particles acquire mass, and the interaction with the higgs boson is described. There is also some discussion regarding the higgs boson production and decay.

2.2 A Brief Summary of the Standard Model

Particle physics describes the world of fundamental particles and the forces between them, depending on the essential theories such as *Quantum Electro-Dynamics QED* which describes mainly the interaction between the light and electrically charged matter. *Quantum Chromo-Dynamics QCD* describes the strong force which is exchanged between elementary particles called quarks and gluons. The general theory which describes QED and QCD called Quantum field theory *QFT*. The third force is the weak force which is mediates some radioactive phenomena and causes transformation of neutron to proton and neutrino ($n \rightarrow p + \bar{e} + \bar{\nu}_e$). Later, QED and the weak force were unified to make a model based on the spontaneously broken direct product group $SU(2) \times U(1)$ called Electro-Weak gauge theory proposed by Glashow, Salam and Weinberg [12]. A more general theory includes both of QCD and electro-weak theory. This is known as the Standard Model which is described in many text books and reviews, see [13–15]. The variety of particles and the interaction between them are organized in the SM. Within the SM, there are two types of fundamental indivisible particles: fermions which have $\frac{1}{2}$-integer spin, and describe

the particles that comprise matter and anti-matter. Fermions have a further classification in to two families. Each family includes six particles which are grouped into three generations. The fermion types are: quarks which carry a charges (color, electric and weak charges) associated with all the fundamental forces, leptons which carry electric and weak charges and they do not participate in the strong force. The second kind of particles is the force mediators or gauge bosons. They have integer spin and carry the three kinds of force between the particles; Electromagnetic *EM*, weak and strong forces. The EM force is mediated by the photon. The heavy electrically charged $W^{\pm}$ and neutral Z bosons carry the weak force. The massless gluon is resposible for the strong force. The Higgs field interacts with the elementary particles and gives them their masses and the minimal excitation state is represented by the massive Higgs boson. Its spin is zero and it has no electric or colour charge [5]. The groupings and general properties of the elementary particles of the SM are summarized in Figure 2.1 and Table (2.1).

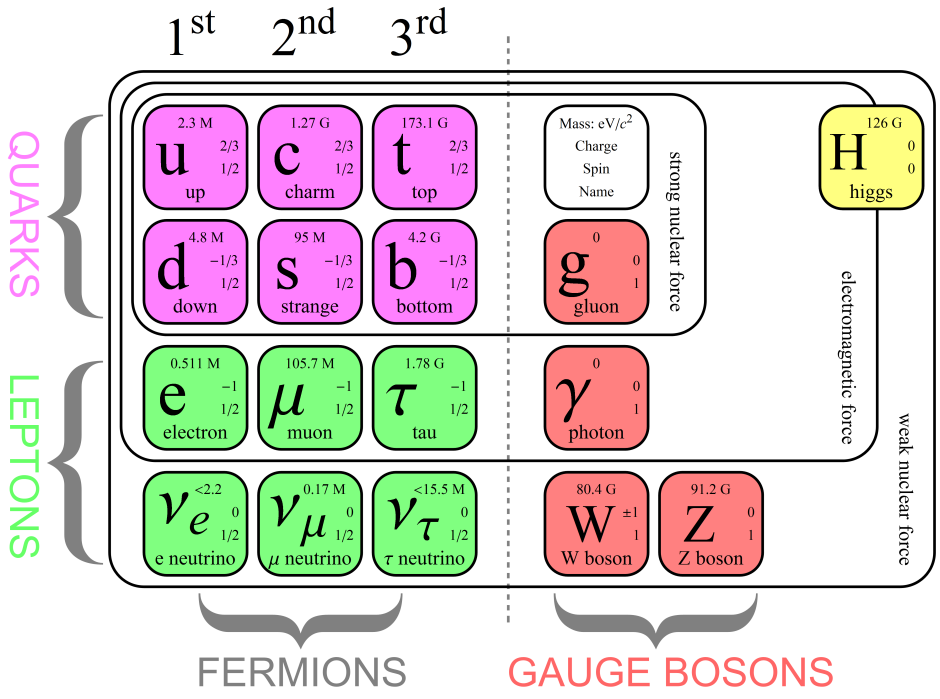


Figure 2.1: The Standard Model of elementary particle physics, with the three generations of fermions which comprise matter. The fourth and fifth columns contain gauge bosons which provide three forces: electromagnetism, the strong nuclear force and the weak nuclear force. The Higgs boson in the sixth column provides an explanation for how the other particles acquire mass

.

Particles	Name	Mass
Fermions	up quark (u)	$2.3^{-0.5}_{+0.7}$ MeV
	down quark (d)	$4.8^{+0.5}_{-0.3}$ MeV
	electron (e)	$0.510998928 \pm 0.000000011$ MeV
	electron neutrino (ν_e)	< 2 eV
	charm quark (c)	1.275 ± 0.025 GeV
	strange quark (s)	95 ± 5 MeV
	muon (μ)	$105.6583715 \pm 0.0000035$ MeV
	muon neutrino (ν_μ)	< 2 eV
	top quark (t)	$173.21 \pm 0.51 \pm 0.71$ GeV (direct measurements)
	bottom quark (b)	4.18 ± 0.03 GeV (MS)
	tau (τ)	1776.86 ± 0.12 MeV
	tau neutrino (ν_τ)	< 2 eV
Gauge Bosons	photon (γ)	$< 1 \times 10^{-18}$ eV
	W	80.385 ± 0.015 GeV
	Z	91.1876 ± 0.0021 GeV
	gluon (g)	0
	Higgs	125.09 ± 0.24 GeV

Table 2.1: The masses of fundamental particle physics [16].

2.3 Gauge Field Theory

Quantum Electro Dynamic field equations can be derived through the use of "gauge field theories" [17]. We often use Lagrangians in Quantum Field Theory (QFT) because they show the particles, masses and interactions. Also, the field equations can be derived from the Lagrangian by minimizing the associated action. By applying the simplest gauge transformation firstly on the Lagrangian of the EM field, we obtain a gauge field, its mass, and an extra term which explains the interaction between electrically charged particles. Moreover, this gauge field theory can be extended into more complex symmetry groups, which explains the experimentally observed force carriers such as γ, Z, W and H.

The Lagrangian of particles obeying the Dirac equation is given by

$$\mathcal{L}_D = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi. \tag{2.1}$$

This Lagrangian is invariant under the group of unitary 1×1 transformations $U(1)$

$$\Psi^{'} = e^{i\alpha}\Psi, \tag{2.2}$$

where α is a global variable which causes a conceptual problem: it is not consistent with special relativity to allow a global phase transformation at all points in space at the same time. Therefore, the parameter is chosen to be a local function of space $\alpha(\bar{x})$. However, the Lagrangian is now not invariant under the transformation

$$\Psi^{'}(x) = e^{i\alpha(\bar{x})}\Psi(x) \tag{2.3}$$

since

$$\partial_\mu \Psi^{'} \neq e^{i\alpha(\bar{x})}\partial_\mu \Psi. \tag{2.4}$$

To circumvent this problem, the partial derivative ∂_μ can be changed to a *covariant derivative* D_μ, which includes the gauge field A_μ

$$D_\mu = \partial_\mu - ieA_\mu, \tag{2.5}$$

where e denotes the electric charge.

Using the covariant derivative D_μ the invariance of the Lagrangian is preserved under the local transformation

$$D_\mu^{'}\Psi^{'} = e^{i\alpha(\bar{x})}D_\mu \Psi. \tag{2.6}$$

By substituting Equation (2.5) into (2.6), we will derive a definition of $A_\mu^{'}$

$$A_\mu^{'} = A_\mu + \frac{1}{e}\partial_\mu \alpha(\bar{x}) \tag{2.7}$$

and the Lagrangian is now invariant under the local transformation Equation (2.3). To have a more complete definition of the Lagrangian, it is important to add a kinetic term for the gauge field A_μ. This term should be quadratic in the first derivative of the gauge field. It must be both Lorentz and gauge invariant. For this purpose, we construct the field strength tensor

$$F_{\mu\nu} = \frac{i}{e}[D_\mu, D_\nu] \tag{2.8}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ie[A_\mu, A_\nu] \tag{2.9}$$

which, in the case of Abelian group (here, $U(1)$) becomes

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.10}$$

Now a more complete definition of the Lagrangian is

$$\mathcal{L}_{(D+EM)} = \bar{\Psi}(i\gamma\partial_\mu - m)\bar{\Psi} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + e\bar{\Psi}\gamma^\mu\Psi A_\mu. \tag{2.11}$$

The constant $-\frac{1}{4}$ indicates that the Lagrangian equations of motion match exactly Maxwell's equations. The local transformation $\alpha(\bar{x})$ has allowed us to introduce the gauge field A_μ which mediates the interaction between the electrically charged particles. Also, the field strength $F_{\mu\nu}F^{\mu\nu}$ is Lorentz and gauge invariant, as required. The third term in Equation 2.11, $e\bar{\Psi}\gamma^\mu\Psi A_\mu$ can be considered as the interaction of the Dirac field with the EM field A_μ.

2.4 Electroweak Interaction

The EM and weak forces can be unified using a direct product of $SU(2)_L \times U(1)_Y$ known as the Glashow-Weinberg-Salam (GWS) symmetry group. The $SU(2)_L$ is the group of 2×2 Special Unitary matrices which represents weak isospin and it affects left-handed fermions (or right-handed antifermions) which participate in the weak interactions. The $U(1)_{EM}$ gauge boson, described in the last section, is a mixture of components from the $SU(2)_L$ and $U(1)_Y$ gauge groups. The $U(1)_Y$ unitary 1×1 matrices represent the weak Hypercharge. Weak Hypercharge is defined as

$$Y = Q - I_3, \tag{2.12}$$

where Q is the electric charge and I_3 is the third component of weak isospin. The $SU(2)$ group is characterized by non-commuting operators known as Pauli matrices $\sigma_i, i = 1, 2, 3$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which together form a set of generators of the $SU(2)$ group elements.

As a consequence, more complex gauge transformations are used

$$\Psi' = U\Psi, \tag{2.13}$$

$$\text{where } U = e^{i\alpha_a(\bar{x})T^a} \tag{2.14}$$

and the T^a are the generators.

The covariant derivative D_μ in Equation (2.5) is rewritten as follows

$$D_\mu = \partial_\mu + igT^a A^a_\mu, \tag{2.15}$$

the arbitrary constant g is the coupling constant.

In the direct product GWS symmetry group of $SU(2)\times U(1)$, the covariant derivative becomes

$$D_\mu = \partial_\mu + igT^a W^a_\mu + ig^{'} Y B_\mu. \tag{2.16}$$

The physical bosons in the previous equation consist of two electrically charged particles $W^\pm_\mu$

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \pm iW^2_\mu), \tag{2.17}$$

and two neutral gauge bosons Z_μ and A_μ, which are represented as a linear combination of W^3_μ and B_μ fields

$$A_\mu = B_\mu cos\theta_W + W^3_\mu sin\theta_W, \tag{2.18}$$

$$Z_\mu = -B_\mu sin\theta_W + W^3_\mu cos\theta_W. \tag{2.19}$$

The angle θ_W is the weak mixing angle, which measures the relative strengths of $SU(2)$ gauge (W_μ) coupling g and U(1) gauge (B_μ) coupling $g^{'}$

$$g^{'}/g = tan\theta_W. \tag{2.20}$$

The weak mixing angle θ_W also gives us the relation between the electromagnetic coupling and the weak and $U(1)_Y$ symmetry coupling through the equation

$$e = g\ sin\theta_W = g^{'} cos\theta_W. \tag{2.21}$$

Thus, the direct product of $SU(2)_L \times U(1)_Y$ will provide us with four gauge fields, $W^1_\mu, W^2_\mu, W^3_\mu$ and B_μ. These fields are found to be massless; the Lagrangian is not invariant under the brute-force introduction of mass terms of the form $\frac{1}{2}m^2 W^a_\mu W^{a\mu}$. However, three of the experimentally observed electroweak gauge bosons are massive $(W^\pm, Z^0)$ and only one is massless (photon γ). To be consistent, it is important to introduce the masses and yet preserve the invariance and renormalizability of the Lagrangian. To accomplish this, the Higgs mechanism is introduced through Spontaneous Symmetry Breaking.

2.5 The Higgs sector and Spontaneous Symmetry Breaking

Spontaneous Symmetry Breaking (SSB) begins by introducing a complex scalar field Φ, and its potential $V(\Phi)$ (see Figure 2.2).

$V(\Phi)$ forms part of the Lagrangian and should preserve the invariance of the Lagrangian. The invariance of the Lagrangian restricts the form of the potential $V(\Phi)$ under the $U(1)$ transformation of $\Phi' = e^{i\theta}\Phi$. We find that terms of the form $\Phi\Phi^*$ and $|\Phi\Phi^*|^2$ are allowed, and we can postulate $V(\Phi)$ of the form

$$V(\Phi) = -\mu^2\Phi\Phi^* + \lambda|\Phi\Phi^*|^2, \tag{2.22}$$

with $\mu^2 > 0$ and $\lambda > 0$. The potential $V(\Phi)$ behaves well at large values of $|\Phi|$ because λ is positive (otherwise the potential will be unbounded). At small values of $|\Phi|$ the behavior of $V(\Phi)$ depends on μ^2, which as mentioned should have a positive value. Therefore, the potential $V(\Phi)$ in Equation (2.22) has a local minimum which occurs at

$$|\Phi_0| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}. \tag{2.23}$$

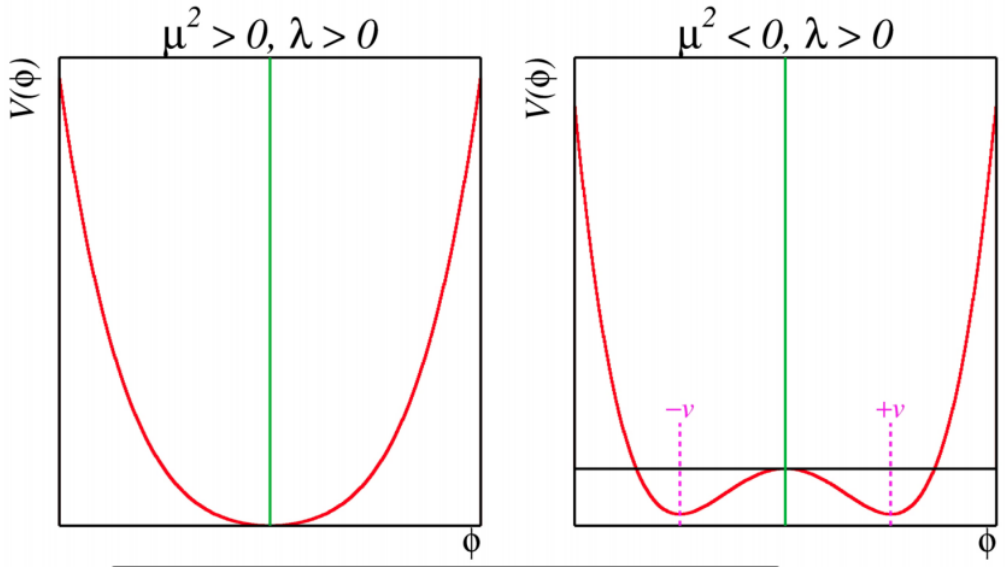


Figure 2.2: The scalar potential $V(\Phi)$. The Higgs field adopts a non-zero (right) vacuum expectation value.

In two dimensions, we can perform a $U(1)$ transformation

$$\Phi^{'} = e^{i\theta}\Phi. \tag{2.24}$$

It is clear from Equation (2.24) that there are many infinite minima at

$$|\Phi_0| = e^{i\theta}\frac{v}{\sqrt{2}} \text{ (where } \theta = 0 \rightarrow 2\pi). \tag{2.25}$$

A specific value of θ can be chosen to be our vacuum, which will break the $U(1)$ symmetry. To simplify our calculations, we will choose $\theta = 0$, which produces a vacuum expectation value *vev* $\langle\Phi\rangle$

$$\langle\Phi\rangle = \frac{v}{\sqrt{2}}. \tag{2.26}$$

The infinite minima allow us to introduce small perturbations around the direction θ and in the radial direction. Consider the perturbations, H in the radial direction, and ϕ/v in the θ-direction

$$|\Phi_0| = \frac{e^{i\phi/v}}{\sqrt{2}}(v + H) \tag{2.27}$$

because $H \ll v$

$$|\Phi_0| \simeq \frac{1}{\sqrt{2}}(\frac{\mu}{\sqrt{\lambda}} + H + i\phi). \tag{2.28}$$

By substituting Equation (2.28) into Equation (2.22)

$$V(\Phi) = \mu^2H^2 + \mu\sqrt{\lambda}(H^3 + \phi^2H) + \frac{\lambda}{4}(H^4 + \phi^4 + 2H^2\phi^2) - \frac{\mu^4}{4\lambda}, \tag{2.29}$$

we can derive from Equation (2.29) the mass term for the H field which is $\mu^2H^2 = \frac{M_H^2}{2}H^2$. So, the mass is

$$M_H = \sqrt{2}\mu. \tag{2.30}$$

It is noteworthy that there is no mass term for the ϕ field which is called the "Goldstone boson". Now, by inserting Equation (2.28) in to the kinetic term $(D_\mu\Phi)^*D^\mu\Phi$, it becomes

$$\begin{aligned}(D_\mu\Phi)^*D^\mu\Phi = \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}g^2v^2A_\mu A^\mu + \frac{1}{2}g^2A_\mu A^\mu(H^2 + \phi^2) \\ - gA_\mu(\phi\partial_\mu H - H\partial_\mu\phi) + gvA_\mu\partial^\mu\phi + g^2vA_\mu A^\mu H.\end{aligned} \tag{2.31}$$

Again, there is no mass term for the ϕ field, which means that the Goldstone boson is massless. However, we do have a mass term for the gauge field A_μ which is $\frac{1}{2}g^2v^2A_\mu A^\mu$ and its mass is

$$M_A = gv. \tag{2.32}$$

Furthermore, the interaction term $g^2 v A_\mu A^\mu H$ gives us the coupling of the gauge field to the H field, which is proportional to the mass of the gauge boson itself

$$g^2 v = g M_A. \tag{2.33}$$

To remove the field ϕ, we must turn from global to local transformations. The choice of the gauge is equivalent to choosing a special direction of the local transformation. In fixing the gauge, the imaginary part of Φ will be removed. Under the gauge transformation we must have

$$\Phi^{'} = e^{-i\phi/v}\Phi, \tag{2.34}$$

where Φ is defined as

$$\Phi^{'} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix}. \tag{2.35}$$

Therefore, the more complete definition of the Higgs-Sector Lagrangian in the unitary gauge is

$$\begin{aligned} \mathcal{L} = \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{M_A^2}{2} A_\mu^{'} A^{'\mu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{M_H^2}{2} H^2 + \frac{g^2}{2} A_\mu^{'} A^{'\mu} H^2 \\ + g M_A A_\mu^{'} A^{'\mu} H - \frac{\lambda}{4} H^4 - \sqrt{\frac{\lambda}{2}} M_H H^3. \end{aligned} \tag{2.36}$$

By applying SSB to the electroweak interaction model $SU(2)_L \times U(1)_Y$, we can obtain a mass for Z and $W^\pm$ bosons, while the photon is still massless.

Firstly, we will examine $D_\mu^{'}\Phi^{'}$

$$D_\mu^{'}\Phi^{'} = (\partial_\mu + ig\frac{1}{2}T^a W_\mu^a + ig^{'}\frac{1}{2} Y B_\mu)\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{2.37}$$

$$= \frac{(v+H)}{\sqrt{2}}[\frac{g}{2}(W_\mu^1 - iW_\mu^2) + (\partial_\mu - i\frac{g}{2}W_\mu^3 + ig^{'}\frac{1}{2}B_\mu)], \tag{2.38}$$

where, $Y(\Phi) = \frac{1}{2}$. The expression for the kinetic part of the Lagrangian is

$$(D_\mu^{'}\Phi^{'})^2 = \frac{1}{2}(\partial_\mu H)^2 + \frac{g^2 v^2}{4} W_\mu^+ W^{\mu-} + \frac{g^2 v^2}{8\ cos^2\theta_W} Z_\mu Z^\mu + (0 A_\mu A^\mu). \tag{2.39}$$

From Equation (2.39) we can define the masses of the photon γ, M_W and M_Z as follows

$$M_W = \frac{1}{2} v g \tag{2.40}$$

$$\begin{aligned} M_Z &= \frac{1}{2}\frac{gv}{cos\theta_W} \\ &= \frac{1}{2} v \sqrt{g^2 + g^{'2}} \end{aligned} \tag{2.41}$$

$$m_\gamma = 0. \tag{2.42}$$

The SM does not predict the values of $g^{'}$ and g and they must be measured experimentally e.g. by using the ratio of M_W/M_Z and Equation (2.21).

For the Higgs boson, its mass is given by Equation (2.30) which can be rewritten as

$$M_H = \sqrt{2\lambda v^2}. \tag{2.43}$$

The value of v is known from the relation between the Fermi constant G_F in weak processes (e.g., week decays of muon) and electroweak coupling

$$\text{muon decay}: \quad \frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{g^2}{8(\frac{1}{2}vg)^2} \tag{2.44}$$

M_W is given by the Equation (2.40) and $G_F = 1.166 \times 10^{-5}$. So, v has the value $v = 246$ GeV [18]. However, λ is a free parameter. Therefore, the mass of Higgs boson cannot be predicted by the SM [19].

2.6 Fermion Masses

It is observed experimentally that the decay of ^{60}Co nuclei and other processes violate parity [20]. Therefore, there is a difference in the behavior of left and right-handed particles in the SM. We express this by using $SU(2)$ doublets for left-handed particles because they experimentally feel the weak force, and singlets for right-handed particles which do not feel the weak force. The left-handed component of all fermions (i.e. quarks and leptons), participates in the weak interaction.

Considering only the first generation of quarks and leptons having left-handed $SU(2)$ doublets and right-handed singlets, we have

$$l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix};\ e_R;\ q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix};\ u_R;\ d_R.$$

To let the fermions interact with the Higgs field, and gain from its mechanism for mass generation, we should deal with them as massless at high energy under $SU(2)$ transformation, and then calculate the mass induced by SSB. In the Dirac Lagrangian the fermion masses arise from the mass term $m\bar{\Psi}\Psi$ and only the terms mixing helicities $m\bar{\Psi}_L\Psi_R,\ m\bar{\Psi}_R\Psi_L$ survive. Therefore, we do not yet have explicit mass terms for the quarks or leptons.

A proposed interaction which generates fermion masses is the "Yukawa interaction" [21]. This couples the scalar Higgs field Φ to the fermionic fields $\bar{\Psi},\ \Psi$. The mass term for the first lepton generation is

$$\mathcal{L}_{lepton\ mass} = -Y_e \bar{l}_L^i \Phi_i e_R + \text{ hermitian conjugate (h.c.)} \tag{2.45}$$

To find these masses we use the same strategy, by promoting the global symmetry to local symmetry and rewrite the Higgs field as

$$\Phi = \frac{e^{i(w_a T^a - BY)}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{2.46}$$

Then, under the gauge transformation, the Goldstone Bosons $w_a,\ B$ will be absorbed.

By introducing this doublet scalar Higgs field, the term which includes the mass and the coupling of the first generation lepton becomes

$$\begin{aligned} \mathcal{L}_{lepton\ mass} &= -\frac{Y_e}{\sqrt{2}} (\bar{\nu}_L \bar{e}_L) \begin{pmatrix} 0 \\ v + H \end{pmatrix} e_R + \text{h.c} \\ &= -\frac{Y_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) v - \frac{Y_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) H \\ &= \frac{Y_e v}{\sqrt{2}} \bar{e}e + \frac{Y_e}{\sqrt{2}} \bar{e}eH \end{aligned} \tag{2.47}$$

The first term in Equation (2.47) is treated as an electron mass which is

$$m_e = \frac{Y_e v}{\sqrt{2}} \tag{2.48}$$

By substitution Equation (2.48) into Equation (2.47) we get:

$$\mathcal{L}_{lepton\ mass} = (m_e \bar{e}e + \frac{m_e}{v} \bar{e}eH) \tag{2.49}$$

There are several features which can be inferred from eqs 2.48 and 2.49. The *vev* of the Higgs field generates the electron mass. Also, the coupling constant Y_e describes the Higgs boson coupling to $\bar{e}e$ which is proportional to m_e. The second term is the fermion coupling to the Higgs boson H which, again is proportional to the electron mass.

For the generations of fermions, we have new Yukawa coupling constants, Y_f. In the quark sector the situation is slightly more complex, where both components of the left-handed doublets acquire mass.

2.7 The Higgs Boson

Given the experimental and theoretical constraints on the expected mass of the Higgs boson, various searches have been executed, leading to the eventual discovery of this particle in July 2012 by the two LHC experiments, ATLAS [22] and CMS [23]. The ATLAS dataset used for discovery corresponds to an integrated luminosity 5.8-5.9 fb^{-1} using pp collisions at a centre-of-mass energy of 8 TeV.

The discovery provides evidence for Higgs boson decays in the following modes $ZZ^* \to 4\ell$, $\gamma\gamma$, $WW^* \to l\nu l\nu$, $\tau\tau$, $\mu\mu$ and bb [24–27]. Then the combination of the Run 1 data from the ATLAS and CMS experiments, leading to improved precision for this new particle mass and confirming that the mass of Higgs boson is 125.09$\pm$0.21(stat.)$\pm$0.11(syst.) GeV [28]. The combination is performed using only the four-lepton invariant mass, $m_{4\ell}$, deriving from the $H \to ZZ^*$ decay channel, and the diphoton invariant mass distribution in the $H \to \gamma\gamma$ decay channel, because these two channels offer the best mass resolution. This Higgs boson mass is consistent with the theoretical an experimental limits discussed above. The Higgs boson decay modes are dependent on its mass and at 125 GeV, the decay to $\bar{b}b$ dominates. However, this is a challenging decay mode because it involves b-tagging the jets with very large backgrounds. If we associate the decay of $H \to \bar{b}b$ with production of $\bar{t}t$, the backgrounds become more manageable. The $t\bar{t}H(H \to b\bar{b})$ channel also allows measurement of top and b-quark Yukawa couplings to the Higgs field. The other processes such as gluon fusion involove a fermion loop, which complicates the extraction of the Yukawa couplings [29]. The fact that the top Yukawa coupling is much stronger than the other quarks couplings, due to the large mass of the top quark, which gives us the chance to perform direct measurement of this coupling, rather than the other Higgs production modes which are only sensitive through loop effects. Moreover, the expected value of the top Yukawa coupling, which is close to unity might be sensitive to new physics [7].

2.7.1 Higgs Boson Production

The cross-section of $pp \to H$ is dominated at the LHC by gg fusion at the Higgs mass around 125 GeV Figure 2.3. The next highest production cross-section is $pp \to qqH$ which includes vector boson fusion, where WW and ZZ bosons fuse to produce the Higgs boson. The third and fourth most important production processes are $pp \to WH$ and $pp \to ZH$ respectively. This is called *associated production* with W and Z bosons. Another form of associated Higgs production is $pp \to ttH$, where

the Higgs boson is produced along with $t\bar{t}$. The dominant leading order Feynman diagram for this process is shown in Figure (2.4) (bottom-left).

In general, with no cleaning cuts applied, the production channel $pp \to H$ is swamped by QCD backgrounds, so it is very hard to find the Higgs boson. The other production modes ($pp \to qqH$, $pp \to WH$, $pp \to ZH$, $pp \to ttH$) are easier to isolate from QCD background by virtue of the fact that there are either two heavy tops or W/Z or extra high-p_T jets to separate them from the QCD background, where heavy particles are not common.

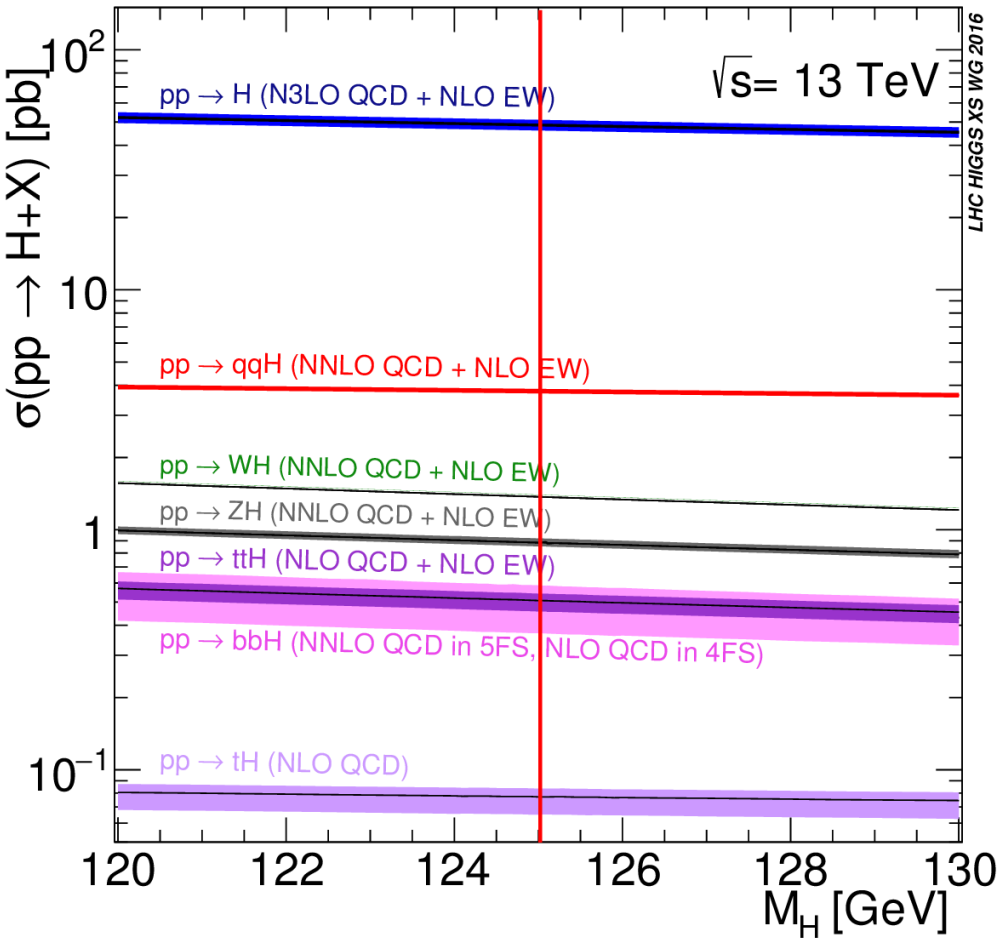


Figure 2.3: Cross section of Higgs production channels [6].

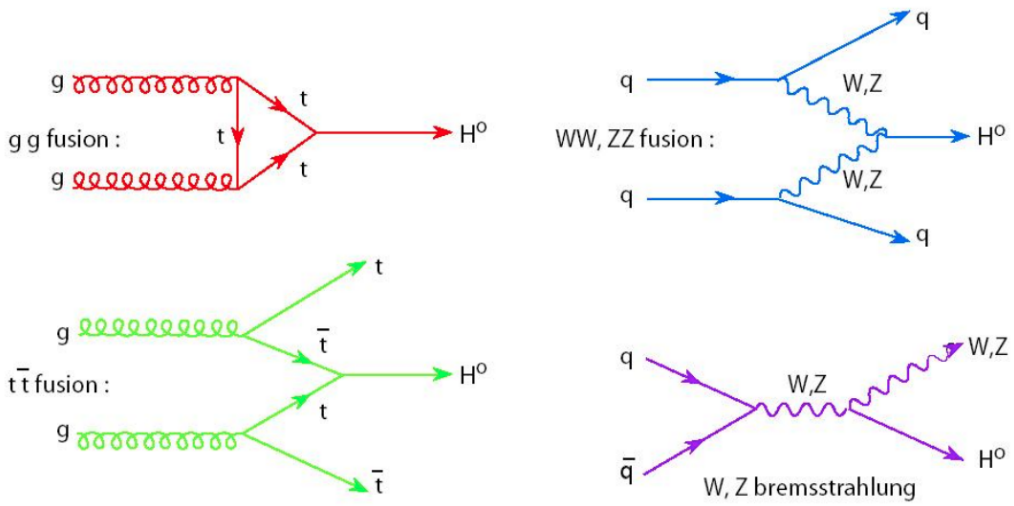


Figure 2.4: Feynman diagrams of Higgs production channels.

2.7.2 Higgs Boson Decay

The main decay of the Higgs boson for masses around 125 GeV is to a $b\bar{b}$ pair, (see Figure 2.5) [30]. This indicates the importance of the $t\bar{t}H(H \to b\bar{b})$ channel, which has some merit in Higgs boson mass measurement, but also in the Yukawa coupling measurements.

By considering the branching ratios of different decay modes of the Higgs boson, we will have a better idea of the physics signatures to search for. The Higgs boson can decay to pairs of electroweak gauge bosons $H \to W^+W^-, ZZ$ or to massive fermion anti-fermion pairs $H \to f\bar{f}$ Figure 2.6. The Higgs can also decay to massless final state bosons e.g. gg and $\gamma\gamma$ pairs via fermion and gauge boson loop diagrams. [31]-[32]. An additional very small branching ratio is the decay $H \to \gamma Z$ Figure 2.5.

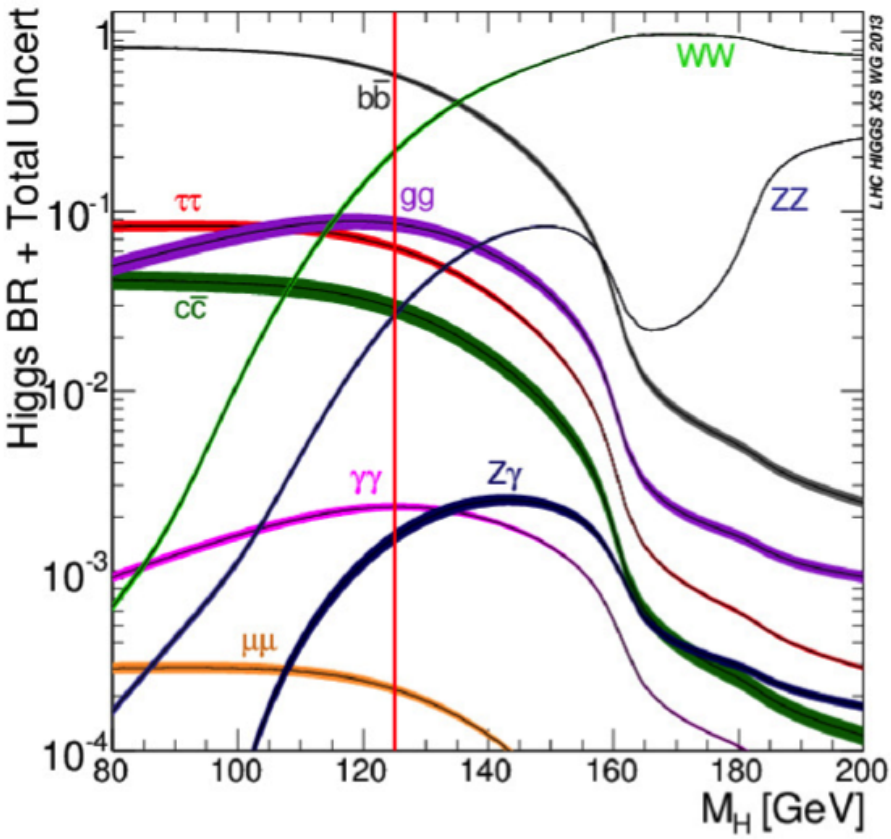


Figure 2.5: Branching ratio of Higgs decay modes [33].

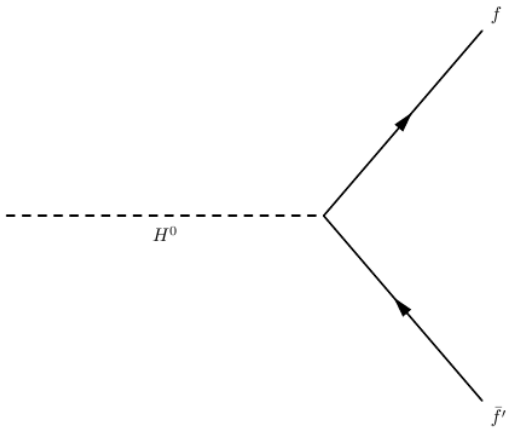


Figure 2.6: Feynman diagram of Higgs decay to fermions.

We derive the Higgs partial decay width into fermions by considering the transition amplitude having matrix elements M

$$\begin{aligned} -iM &= \bar{u}^{s_1}(p_1)(im_f/v)v^{s_2}(-p_2) \\ +iM &= \bar{v}^{s_2}(-p_2)(-im_f/v)u^{s_1}(p_1) \end{aligned} \tag{2.50}$$

The squared transition amplitude is:

$$\begin{aligned}\sum |M|^2 = (\frac{m_f}{v})^2 N_c \mathrm{Tr}(\not{p}_1 + m)(-\not{p}_2 - m) \\ = 2N_c(\frac{m_f^2}{v^2}) M_H^2 \beta_f^2\end{aligned} \tag{2.51}$$

Where N_c is the number of colors (1 for leptons and 3 for quarks) and $\beta_f = (1 - \frac{4m_f^2}{m_H^2})^{\frac{1}{2}}$ is the velocity of the fermions in the final state.

So, the partial decay width is:

$$\Gamma(H \to f\bar{f}) = N_c \frac{m_f^2}{v^2} \frac{M_H}{8\pi} \beta_f^3 \tag{2.52}$$

The decay width grow with M_H. Figure 2.7 shows the sum over all kinds of Higgs boson decay modes.

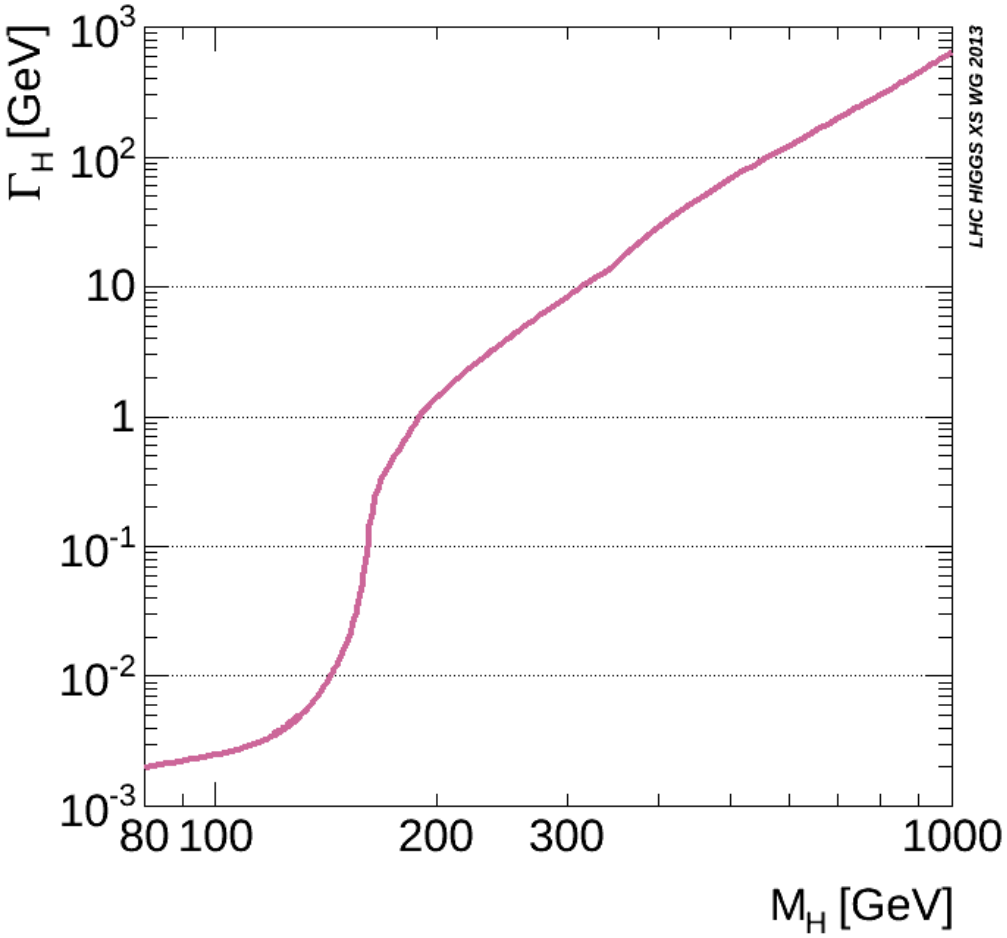


Figure 2.7: Total decay width of Higgs boson decay modes [33].

It is clear that for a hypothetical Higgs boson with mass around twice the W mass, the decay of the Higgs switches predominantly to WW bosons. This causes the decay

width to become much larger. So, a low mass Higgs boson will have a small width and the discovery of the Higgs boson having a mass of 125 GeV means the total decay width is dominated (57.8%) by the decay $H \to b\bar{b}$. However, the relatively high number of background events in gluon fusion with $H \to b\bar{b}$ compared with number of signal events makes such a search unfeasible in any case. When we consider the decay $H \to b\bar{b}$ in association with $t\bar{t}$ pair, the channel becomes less difficult but still very challenging [30]. As a consequence, to observe the $H \to b\bar{b}$ decay, it is important to reconstruct the different $t\bar{t}H$ signatures which are distinguished by type of top quark decay mode.

The top quark decay probability of t$\to Wb$ is almost 100% (99.8 %) [16] and it decays immediately before the hadronization can take place. Then, the final state of $t\bar{t}$ can be divided into three classes according to the W boson decay. The first type is the di-leptonic channel: $t\bar{t} \to \bar{l}\nu_l b l^{'} \bar{\nu_{l'}} \bar{b}$. The second type is the semi-leptonic channel: $t\bar{t} \to \bar{l}\nu_l b q \bar{q} \bar{b}$. Finally, there is the all-hadronic channel: $t\bar{t} \to q\bar{q}bq^{'}\bar{q}^{'}\bar{b}$. The W boson decays around 2/3 of the time into quark-antiquark pairs and 1/3 of time into lepton-neutrino pair. The following Table 2.2 displays the measured BR of different W^+ decay modes.

Decay mode	Branching Ratio (%)
$W^+ \to e^+\nu_e$	10.71 ± 0.16
$W^+ \to \mu^+\nu_\mu$	10.63 ± 0.15
$W^+ \to \tau^+\nu_\tau$	11.38 ± 0.21
$W^+ \to$hadrons	67.41 ± 0.27

Table 2.2: Branching ratio of W^+ boson decay [16].

The di-leptonic final state has the smallest background whilst it also has the smallest BR (around 10%), see Figure 2.9. In contrast, the all-hadronic final state has a huge background from the QCD multijets production and the largest BR with about 30% (taus are not considered in this study). The last channel which this thesis will investigate is the semi-leptonic final state specifically e+jets and muon+jets Figure 2.8.

The semi-leptonic channel has a large BR with about 30% Figure 2.9 and its relative background ratio is small in comparison with the all-hadronic channel. This property

makes the semi-leptonic final state both a clean signature to trigger on, and provides a good discrimination between signal and background [34].

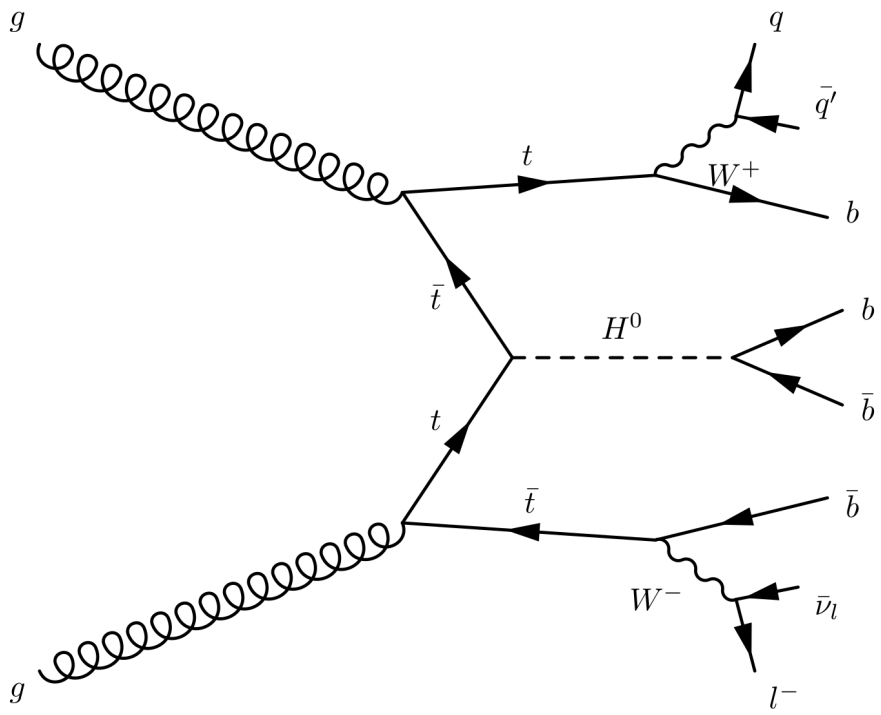


Figure 2.8: Feynman diagram of semi-leptonic decay mode.

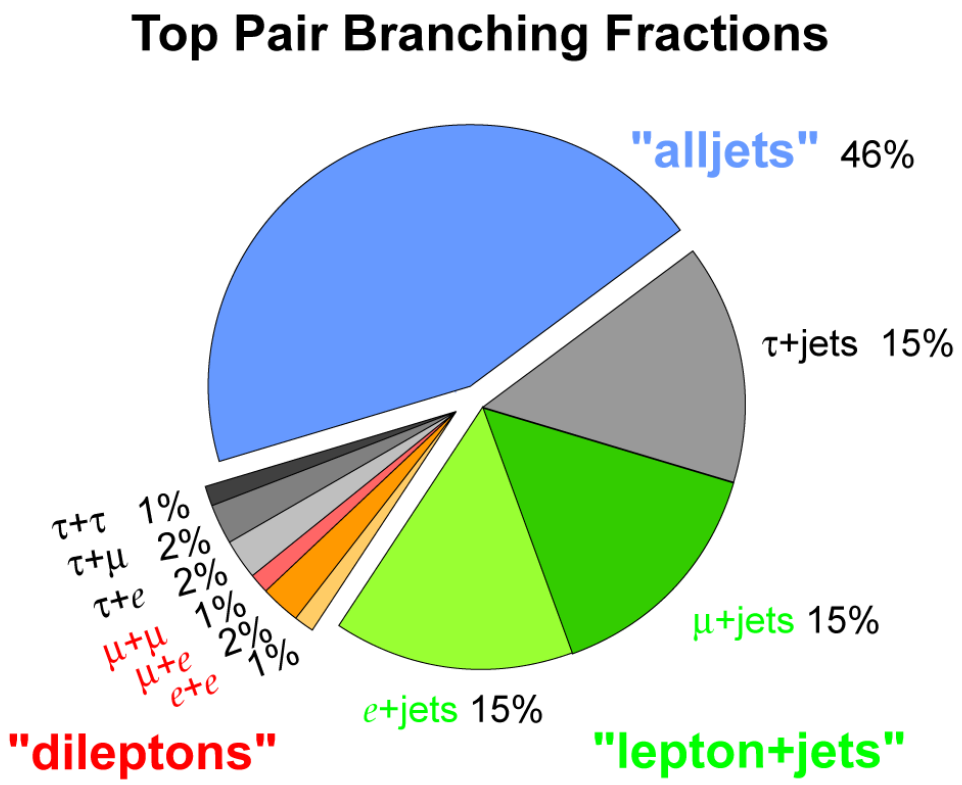


Figure 2.9: Decay modes of the $t\bar{t}$ events classified according to the decay of the produced W-bosons. Neutrinos are not mentioned in the decay labels.

In this Chapter gauge theory is reviewed, along with the Higgs mechanism using spontaneous symmetry breaking to illustrate the process by which the Higgs boson is produced, and also how it decays in various modes at the LHC. Among the most interesting modes is the ttbarH production with associated decay of the Higgs boson to bbbar.

2.8 Summary

In this Chapter gauge theory is reviewed, along with the Higgs mechanism using spontaneous symmetry breaking to illustrate the process by which the Higgs boson is produced, and also how it decays in various modes at the LHC. Among the most interesting modes is the $t\bar{t}H$ production with associated decay of the Higgs boson to $b\bar{b}$.

Chapter 3

The ATLAS Detector

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [35] is the only accelerator to collide protons at a centre-of-mass energy of 13 TeV. It contains two rings of superconducting magnets around the collider ring which lies in a 27 km tunnel. The same tunnel was formerly used for the LEP accelerator [36]. It is located across the border of France and Switzerland at a depth of about 100 m and situated at the Conseil Européen pour la Recherche Nucléaire (CERN). The cern accelerator components are illustrated in Figure 3.1: its ring has eight independent sectors (octants), eight arcs and eight straight sections called *insertions*. The arcs are the location of the magnet system, with its different types, such as the dipole bending magnets, quadrupole magnets and the other multipoles magnets. The insertions, depend on their function, such as usage in injection, Radio Frequency (RF) and beam cleaning. Four of the "insertions" provide location for the LHC detectors.

Two general purpose detectors are located at opposite straight sections. A Toroidal LHC Apparatus (ATLAS) [37] is based at Point 1, and Compact Muon Solenoid (CMS) [38] at Point 5. The other two detectors, A Large Ion Collider Experiment (ALICE) [39] and the Large Hadron Collider beauty (LHCb) [40], are adjacent to ATLAS, at Points 2 and 8 respectively. These detectors are located at the crossing points of the beams.

The data used in this study result from the analysis of the high energy proton-proton collisions in the ATLAS detector.

To reach this high energy, the protons are accelerated in several stages. The protons

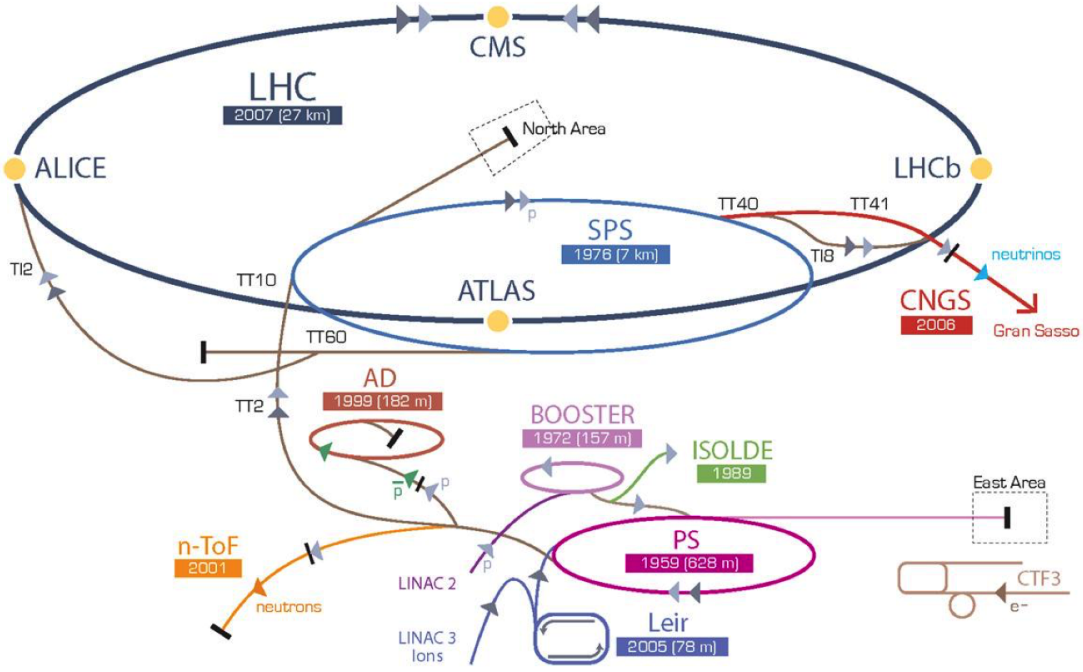


Figure 3.1: Illustrative layout of the LHC

that make the final high energy collisions transverse through an injection chain. The chain consist of a series of machines that accelerate gradually until they reach the required highest energy of 13 TeV as shown in Figure 3.1. The electrons are stripped from their hydrogen atoms by using of an electric field and then they are formed to bunches of protons. These bunches are injected into Linear Accelerator 2 (LNAC2). Then, the Proton Synchroton Booster (PSB) increases the proton energy up to 1.4 GeV. The protons from the PSB are accelerated up to 26 GeV in the Proton Synchroton (PS) ring. Afterwards, the protons are transferred to the underground Super Proton Synchroton (SPS) which is suited for accelerate the protons to 450 GeV before they are injected into the final ring which is LHC. In the LHC, the protons in the beams circulate for many hours and brought into collision at a center-of-mass energy of 13 TeV in ATLAS (in 2015). The protons traverse through injection chain placed at CERN laboratory. Through these different rings, RF is used to accelerate the beam (RF cavities), or to split it into bunches. The injection chain aims to deliver the highest possible proton currents to the LHC and the proton beams should have tightly focused energy and spatial parameters.In order to accomplish this, low energy protons are accelerated in a series of steps using small circularrings which increase in size and energy. The RF cavities in each successive stage add energy and the magnets maintain the beam in a circular trajectory whilst the energy is increased. In the LHC, 1232 superconducting dipole magnets in the arcs are used to steer the beams in circular paths. In order to

focus the beam, quadrupole magnets are used. Quadrupoles known as "QF" focus the beam in horizontal plane and defocus the beam in vertical plan. A second set of quadrupole magnets known as "QD" focus the beam in the vertical plane and defocus the beam in horizontal plane as illustrated in Figure 3.2. In total, the LHC

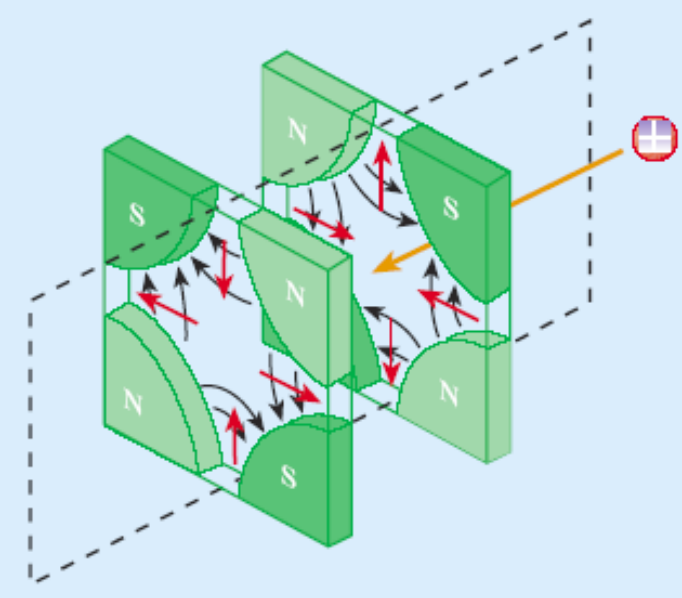


Figure 3.2: Schematic showing horizontal and vertical focusing (QF and QD respectively) quadrupole magnets as used in the LHC.

has 392 quadrupole in both insertion and arc parts. There are also a small number of sextupole, octupole and decapole magnets. The LHC has eight "inner triplet" quadrupole magnets which provide the final focusing of the incident beams at the interaction points [41]. After the injection into the LHC, the dipole magnetic field is increased up to 8.3 T using a current of 11,080 amperes in the superconducting coils which produce the magnetic field. To preserve the superconductivity of the magnets, around 96 tonnes of superfluid helium is cooled down to 1.9 K. Finally, to avoid collisions with gas molecules, the beams are housed in a beam pipe where there is a very strong vacuum.

During the period 2010 to 2013, the first run (RunI) was performed at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. Then, the machine and its experiments upgraded to produce a higher energy and luminosity. The LHC started running again (RunII) in 2015 and it reached an energy of $\sqrt{s} = 13$ TeV which comprises the data used in this thesis.

The beam energy and instantaneous luminosity are the most important parameters in the LHC to quantify the performance of the particle collider. The quantity that measures the ability of a particle accelerator to produce the required number of

interactions in a given time is the instantaneous luminosity. The basic definition of the instantaneous luminosity is the relationship between the number of events per second dN/dt and the cross-section σ_P. The instantaneous luminosity (ℓ) is the proportional factor between them:

$$\frac{dN}{dt} = \ell \cdot \sigma_P \tag{3.1}$$

From the previous relation, the unit of ℓ is Length^{-2}Time^{-1} (σ_P can be expressed in cm^2, 1 barn=10^{-24}cm, and time is chosen to be in seconds). By assuming the beam to have a Gaussian distribution inside the bunches, the ℓ can be written as

$$\ell = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

where $N_{1,2}$ are the number of protons in beam 1,2 respectively, f is the bunch frequency,

N_b is the number of bunches and $\sigma_{x,y}$ are the Gaussian width which is also called the beam size [42]. However, it is important to take into account the effect of the crossing angle to avoid parasitic collisions. This leads to a luminosity reduction by a factor S

$$\ell = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \cdot S \tag{3.3}$$

Luminosity can be also calculated using alternative beam size parameters which have a nominal (i.e. design specification) values in LHC. These parameters are the normalized emittance ϵ_n which is constant a long the ring trajectory and has the nominal value $\epsilon_n = 3.0\mu m$ (in variance of the transverse emittance which is varying along the beam pipe, $\epsilon_n = \beta\gamma\epsilon$ that the β and γ are respectively the velocity of the particle divided by the light velocity and the Lorentz factor) and betatron function (betatron amplitude)$\beta(s)$ which varies along the beam trajectory s and it is a periodic function. It can be interpreted as the local wavelength of the betatron oscillations divided by 2π. The betatron function at the interaction point (IP) (a distance of doubling of width from the interaction point) has a nominal value and it is denoted by β^* =80 cm. So, the luminosity can be written as:

$$\ell = \frac{N_1 N_2 N_B f \gamma}{4\pi\epsilon_n\beta^*}, \quad \text{where } \sigma = \sqrt{\epsilon \cdot \beta^*} \tag{3.4}$$

The LHC design luminosity is $\ell = 10^{34}$ cm$^{-2}s^{-1}$ at 14 TeV [43]. The other factor related to the luminosity ℓ and is essential in ATLAS operation and optimization is

the integrated luminosity over a period of time $\mathcal{L}$ which is defined as

$$\mathcal{L} = \int \ell \mathrm{d}t. \tag{3.5}$$

The integrated luminosity target in ATLAS and CMS is 3000 fb^{-}1 by 2030 [44].

3.2 The ATLAS Detector

The deep understanding of the basic components of matter which are the elementary particles, needs the high energy collisions of particle physics provided by the LHC. This kind of collision provides a wide signature range and unequaled energies and instantaneous luminosity (compared to any other collider) which are studied by the ATLAS detector. ATLAS is one of the two largest general purpose detectors housed at the LHC. It is characterized by roughly full coverage of the solid angle around one of the *pp* interaction points and has a cylindrical geometry.

ATLAS measures about 45m long, 25m in diameter and weighs 7000 tons. This design provides a space for a large magnet system, a toroidal magnet which contains eight independent identical superconducting barrel coils and with two end-caps housing eight further coils each. ATLAS takes its name from the toroidal magnet. In addition to this toroidal magnet field, ATLAS has a solenoid magnet. Both toroidal and solenoidal magnetic fields bend the electrically charged particles' tracks due to the Lorentz force experienced whilst traversing the magnetic field. These magnet systems surround different layers of sub-detector systems. The three main sub-detector components are, the inner Detector (ID) submerged in the solenoidal field which constitutes a tracking system, the calorimeters which are divided into two sub-layers, an Electromagnetic calorimeter (ECAL) and a Hadronic Calorimeter (HCAL), and the outer-most system is the Muon Spectrometer (MS).

The ATLAS Coordinate System

The Cartesian right-handed coordinate system shown diagrammatically in Figure 3.3, is the common coordinate system that is used in ATLAS. The nominal interaction point is chosen as the center of the detector reference system. The x-axis is specified as the direction from the interaction point to the center of the LHC ring; the y-axis is perpendicular to the beam axis and z-axis and it points in the upwards direction; and the z-axis is considered as the counter-clockwise beam direction. In order to acquire an accurate description of the particles' distribution and the efficiency of reconstruction, spherical polar coordinates of the form (r, θ, ϕ)

are used. The azimuthal angle $\phi = tan^{-1}(y/x)$ lies on the $x-y$ plane measuring the space around the beam axis z, the polar angle $\theta = cos^{-1}(z/r)$ is measured in the $r-z$-plane with $r = \sqrt{x^2+y^2+z^2}$. The polar angle is important in defining one of the kinematical variables to describe particles in an event. This variable is pseudorapidity η, (p_T, η, ϕ, E). Hence for relativistic particles, it is easier to measure a particle's θ to calculate η and then to estimate the rapidity y ($\eta \simeq y$)

$$\eta = -ln(tan(\frac{\theta}{2})) \tag{3.6}$$

rather than calculate the rapidity of the particle

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z} \tag{3.7}$$

which need its energy and momentum. [45] [46].

ATLAS has a particle momentum system covering the central range $|\eta| < 4.9$. Since this analysis relies heavily on tracking via the Inner Detector, the pseudorapidity is considered in the region $|\eta| < 2.5$.

Finally, the distance between two particles ΔR is defined in $\eta - \phi$ space as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{3.8}$$

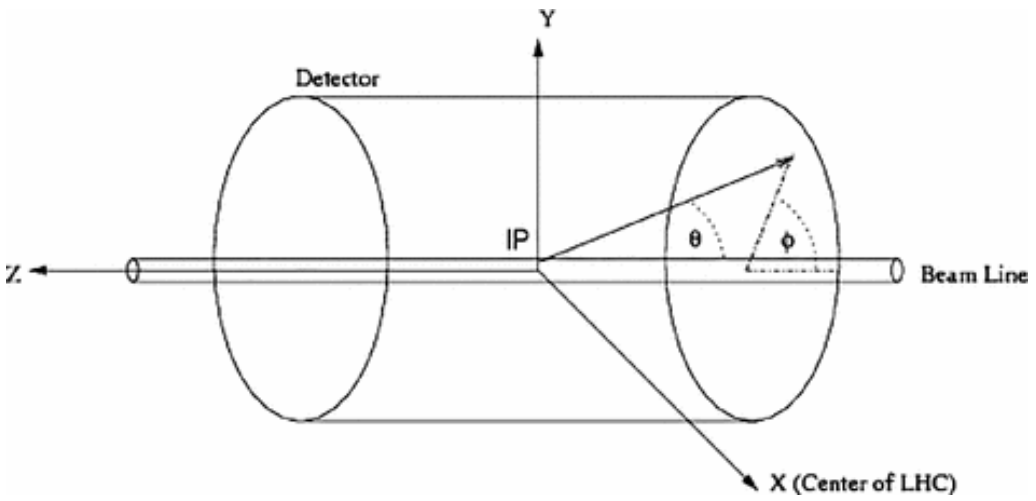


Figure 3.3: Illustration of the ATLAS detector coordinate system.

3.2.1 The Inner Detector

The physics processes resulting from the collisions in the beam require high-precision measurements to achieve a very fine identification of the production particles. Different components are used to recognize the origin of the charged particles, such as

track direction, charge, momentum, impact parameter and the vertices (interaction point of the produced particle or displaced distance from the origin point as the secondary vertex). These components are measured in the Inner Detector (ID) [47] which is embedded in a solenoidal magnet that generates a magnetic field of 2 Tesla with a peak of 2.6 T on the superconductor to obtain a high precision measurement of the transverse momentum [48]. The pseudorapidity region of $|\eta| < 2.5$ is covered by the three sub-components of the ID, the Pixel Detector (Pixel), the Semiconductor Tracker (SCT) [49] and the Transition Radiation Tracker (TRT) [50]. Each sub-detector consists of a barrel part and two end-caps. The layout of the ID and its barrel and end-cap sub-components are displayed in Figure 3.4 .

The ATLAS Pixel detector for LHC RunII consists of four pixel layers in the Barrel region, which increased from three layers to four layers with the addition of the *Insertable B-Layer* (IBL) [52] between RunI and RunI in 2014. The IBL becomes the innermost layer of the Pixel detector (33 mm from the beam axis), which requires a smaller beam pipe radius near the interaction point. It is characterized by higher resolution pixel than the other three layers, to increase the efficiency of reconstruction of the primary and secondary vertices which is essential for the detection of long-lived particles like b-quarks. After insurtion, almost 5 times better rejection against light-jets is achieved in the b-jet tagging algorithm. An improvement by a factor of two in the impact parameter resolution was also observed for low p_T tracks [53]. The IBL adds an extra 12M pixel channel to the former pixel detector, which was 80M channels. The pixel size of the sensors is 50 $\times$ 250 μm^2 in the azimuth- and z-directions, which is 60% of the size of the old pixels (50 $\times$ 400 μm^2). Two different sensor technologies are used; the planar sensor used in 75% of the total IBL volume and 3D sensor technology in the rest.. The second layer in the barrel part is separated from the beam by 50.5 mm. There are three end-cap disks which are installed perpendicular to the beam pipe on each side of the barrel. The pixel detector consists of 1,744 modules each module contains 16 front-end chips (FE) which are equipped with silicon sensors. Each pixel measures 400$\times$50 μm^2 except for a small minority measuring 600$\times$50 μm^2 between the FE chips to achieve the highest granularity. The pixel detector covers the full ID pseudorapidity region of $|\eta| < 2.5$. A particle originating at the interaction point could pass through the barrel crossing at most three pixel layers leaving a hit in each layer. This enables a high precision measurement of he charged particle tracks.

The SCT detector works in a similar way to the Pixel detector, but with long and narrow silicon strips rather than small Pixels to cover a larger surface area

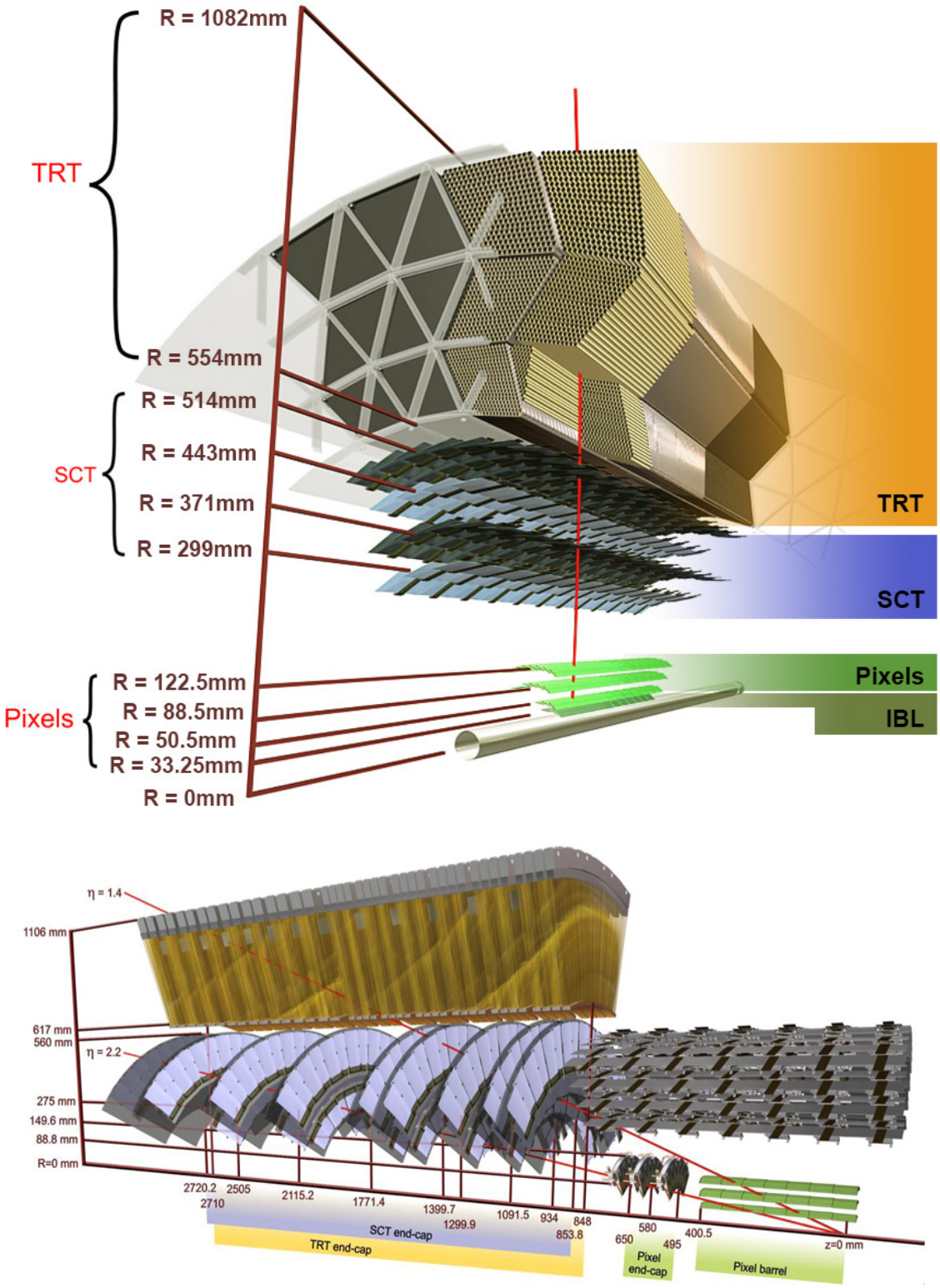


Figure 3.4: The barrel (upper Figure) and the end-cap (lower Figure) scheme of the ATLAS ID [51].

and capture the charged particle tracks. The SCT is the middle part of the ID and it surrounds the Pixel detector. It consists of 4088 modules in total. The 2112 modules are arranged in four double-sided layers in the barrel part, and the remaining modules are distributed among the nine discs in each end-cap to cover the range $|\eta| < 2.5$. Each module consists of two different back-to-back micro-strips with a relative angle of 40 μrad to allow the z-coordinate measurement, and electronics to read-out the signal of the particles (current) on the strips at the top and the bottom of the module. In the SCT, the strips have a pitch of 80 μm which gives a spatial resolution (covering $rcos\theta - \phi$ space) of 17 μm. A resolution of 580 μm in the z-direction is achieved by having the back-to-back modules at a relative angle of 40 μrad. The silicon strips are perpendicular to the beam in the discs of the end-cap and provide 17 μm of spatial resolution in the $rsin\theta - \phi$ space.

The outermost subsystem of the ID is the Transition Radiation Tracker (TRT). The measurements of tracks in the TRT are made in $rcos\theta - \phi$ plane by covering $|\eta| < 2.0$ and gathering on average 36 hits per track, with lower precision than the SCT and Pixel. The TRT modules are a gaseous type of detectors comprised of long cylindrical component called straws which are 4 mm in diameter. The straws are 144 cm long in the barrel (52,544 tubes), and 37 cm long in the end-cap region (245,760 tubes). Each straw has a wire in its center of diameter 31 μm which works as an anode and it is readout individually. The straws are filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2 [54] [55]. The straws are placed to be parallel to z-axis in three rings in the barrel section and radially in the end-caps. The TRT operation depends on two principles [56]; the transition radiation principle (for ATLAS X-ray radiation 10-30 keV) and the ionization of the gas-filled straws by the charged particles energy passing through. Electrons passing through two mediums of different dielectric constant, the polypropylene-polyethylene fibers between the straws and gas-filled straws, produce transition radiation photons which ionize the gas and cause a signal in the anode wire. Electrons produce a larger transition radiation, so it is possible to distinguish between $e^{\pm}$ and other heavy particles. The deposited energies due to transition radiation and ionization are added to calculate the output signal which is the main factor to specify a straw hit.

3.2.2 Calorimeters

The calorimeter system is designed to absorb and measure the energy of different types of particles deposited via electromagnetic or hadronic interaction with the calorimeter material and covers full ϕ space and the range $|\eta| < 4.9$. There are two

different kind of calorimeter in ATLAS; electromagnetic and hadronic calorimeter, see Figure 3.5. Both calorimeter types are designed based on a sampling principle. Two mediums are used in the sampling calorimeter, an absorbing material (passive part) to force the particles to interact with the nuclei of the dense absorber material. This initiates a shower of particles, which will extend to the second (active) material. Particle energies are measured with good precision and fine segmentation in the ATLAS electromagnetic calorimeter and coarser segmentation less precision in the hadronic calorimeter. The difference in the precision of the energy measurement in the two kinds of calorimeter is due to the difference of the nature in interaction with the material nuclei. When the electron meets the absorbers, it interacts electromagnetically with the absorber material and creates a shower of low energy particles. Firstly, an electron radiates a photon γ (Bremsstrahlung) which later produces a pair of e^+e^-. The interaction iterates until the energy of the new particles is below the critical energy E_c. The particle shower resulting from that interaction has a characteristic distance X_0 called *radiation length* and it is fixed by the detector material (it depends on how close the material atoms are and how many charges there are in the nucleus). X_0 in the electromagnetic calorimeter is small than the *nuclear interaction length* λ_I in the hadronic calorimeter.

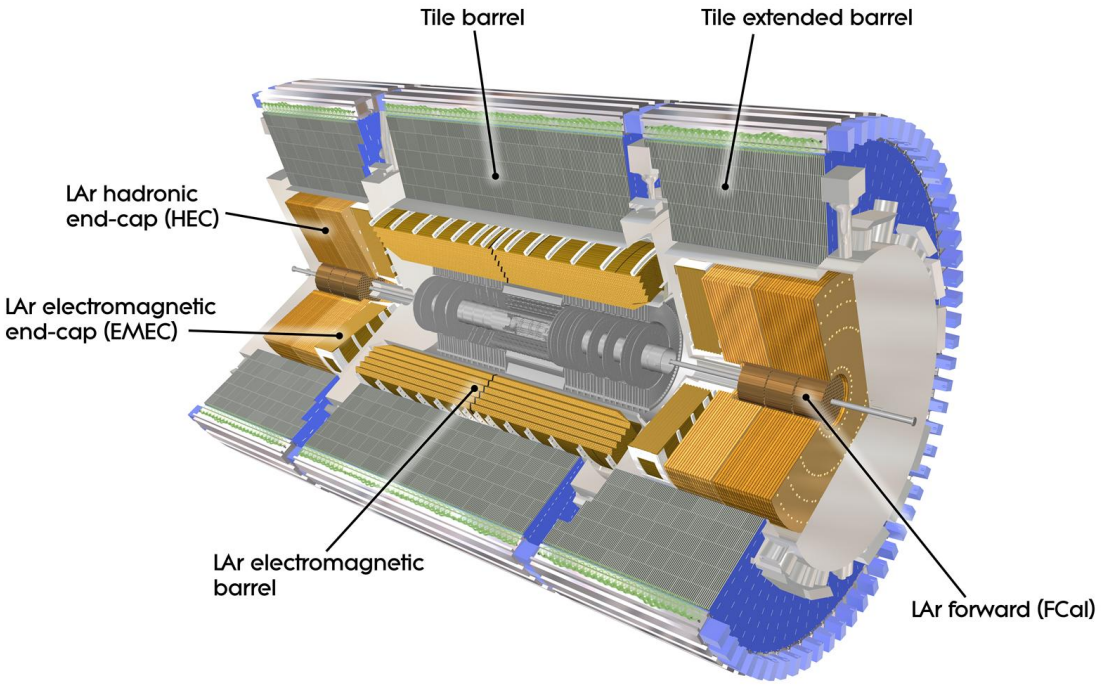


Figure 3.5: The Atlas electromagnetic and hadronic calorimeters in barrel and end-cap region [48].

The statistical uncertainty in the shower energy σ_E eq.(3.10) is proportional to the statistical fluctuations of the average number of charged particles produced N. Typically, the number of particles produced from a signal charged incident particle will be roughly equal to the original particle energy (E) divided by the threshold critical energy E_c

$$N = \frac{E}{E_c}, \tag{3.9}$$

$$\sigma_E \propto \sqrt{N}, \tag{3.10}$$

and the overall energy proportional also with $\sqrt{N}$

$$\sqrt{E} \propto \sqrt{N}. \tag{3.11}$$

So, from eq.(3.10) and (3.11), one can obtain the relative resolution as following

$$\frac{\sigma_E}{E} \propto \frac{1}{\sqrt{E}}. \tag{3.12}$$

and the full energy resolution of the calorimeter can be defined as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} + b + \frac{c}{E}, \tag{3.13}$$

where a is called the stochastic term. The precise value of this constant depends on the calibration of the particles showering giving an energy error proportional to $\sqrt{E}$. It is also the function of the geometry of the detector. The calorimeter precision is a compromise between cost, quality and quantity of sampling which the calorimiter is required to have. The constant term b is due to inhomogeneities across the calorimeter producing errors in the different cells' calibration. The noise term $\frac{c}{E}$ depends on the instantaneous luminosity which includes the pileup. At low luminosity, the electronic noise in the calorimeter can be significant. However, at high energy, the term $\frac{c}{E}$ is usually neglected.

Electromagnetic Calorimeter

The ECAL is used to measure the energy of electrons and photons. In ATLAS the ECAL has an accordion-shaped structure which provides full coverage in ϕ, consists of many layers of lead covered by stainless steel which represents the particle absorber part (passive part). Between them is the read-out part of the calorimeter (electrodes) which consists of three conductive copper layers separated by insulating polyimide sheets. The liquid Argon (LAr), fills the space between electrodes, cooled

to 87K and comprises as the active material because of its radiation hardness and stable response.

The barrel section covers the range of $|\eta| < 1.475$ with its two parts that are separated by a 4mm gap. The radiation length in the barrel ranges from 22 X_0 →33 X_0. The two end-cap sections cover a larger η space, which ranges between $1.385 < |\eta| < 3.2$. The end-caps have a minimum depth of 24 X_0 and a maximum of 38 X_0. The energy resolution of the ECAL can be parameterized as:

$$\frac{\sigma_E}{E} = \frac{(10.1 \pm 0.4)\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus (0.2 \pm 0.1)\% \tag{3.14}$$

The small value of the stochastic term results from the sampling fluctuations.

Hadronic Calorimeter

The HCAL is built immediately outside the ECAL. It measures the energy contained by jets of particles originating from quarks and gluons as well as from the hadronic decays of taus. As is mentioned in the first part of this section, the hadronic showers are longer and wider than electromagnetic ones. Hence, that leads to the requirement for a 'thicker' hadronic calorimeter. The hadronic showering essentially starts in the electromagnetic calorimeter and continues to the hadronic calorimeter where it becomes broader and deposits most of its energy there, according to the length λ_I. The HCAL has a thickness, transverse to beam of around 10 λ_I at minimum in order to absorb the energy of almost all of the particles passing through the calorimeter. The hadronic calorimeter section with its three cylinders in the barrel covers a larger η range than the ECAL. The center part called the Long Barrel (LB) and is split into two sections LBA and LBC. It is a tile calorimeter which covers up to $|\eta| < 1.0$. At each end of the LB are two Extended Barrel (EB's) covering the range $0.8 < |\eta| < 1.7$. The tiles are layers of passive steel and a scintillator as an active material arranged radially to the beam in 256 wedge-shaped segments arranged in ϕ around the beam axis as illustrated in Figure 3.6. The scintillator is coupled to wavelength-shifting fibres which carry the radiation light originating from the absorber nuclei reaction (after the interaction with the incident particles) and is finally readout by photomultipliers [57].

In addition to the hadronic barrel calorimeters (LB, EB's), ATLAS also has two hadronic end-cap calorimeters. These are composed of two cylindrical wheels covering arrange of $1.5 < |\eta| < 3.2$, while the forward calorimeter ECAL covers a higher $|\eta|$ in the range $3.1 < |\eta| < 4.9$. Both end-cap and forward calorimeters use LAr as an active medium but each of them uses a different absorber, copper and tungsten

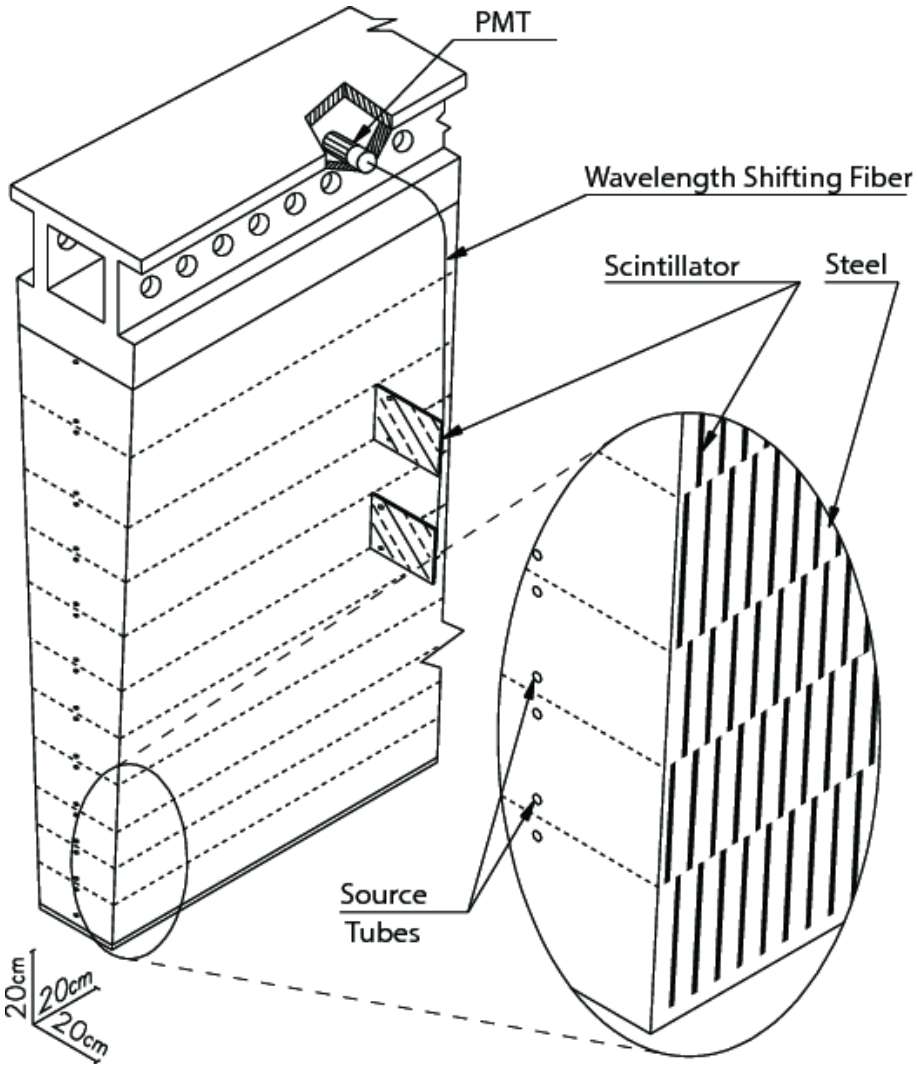


Figure 3.6: Schematic displaying the wedge-shaped module of the HCAL tile barrel [48].

in end-cap and forward calorimeters, respectively. The resolution of the hadronic calorimeter was measured using of pion test beams. The measured energy resolution for the tile calorimeter is

$$\frac{\sigma_E}{E} = \frac{(56.4 \pm 0.4)\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus (5.5 \pm 0.1)\% \tag{3.15}$$

and in the end-cap calorimeter it is

$$\frac{\sigma_E}{E} = \frac{(70.6 \pm 1.5)\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus (5.8 \pm 0.2)\% \ . \tag{3.16}$$

The reason for the higher value of the constant a, see eq.3.13, in the HCAL resolution is that the hadronic shower consists of two distinctly different interactions. There is an electromagnetic interaction for some of the particles coming from the absorption process (e.g pions and muons). Then, there is the strong interaction of everything with absorber material. Here, there is around 40% "invisible" energy due to the release of binding energy from the nuclei of the absorber (nuclear break-up) and from escape of neutrinos [58].

3.2.3 Muon Spectrometer

Muons are one of the ttH channel components in the semileptonic final state. Typically, muons are the only visible particles escaping the calorimeters. Consequently, it is required that the Muon Spectrometer (MS) is situated in the outer most part of the ATLAS detector. The MS is comprised of barrel chambers and end-cap discs. The barrel chambers are arranged in three concentric layers at radii of 5, 7.5, and 10 m from the beam axis covering the central region $\eta < 1$. The η-range has been extended by the end-cap, each containing four wheels at 7.4, 10.8, 14 and 21.5 m from the interaction point (IP), with their centers at the beam axis. The end-caps provides coverage in the range $1.0 < |\eta| < 2.7$. The three dimensional view of the spectrometer is displayed in Figure 3.7. The MS integrates two concepts; "precision chambers" for tracking and "trigger chambers" for timing. The precision tracking is provided by two types of chamber, Monitored Drift Tube (MDT) and Cathode Strip Chambers (CSC) at high $|\eta|$, close to the IP. There are also two types of trigger chamber, Resistive Plate Chambers (RPC) in the barrel region, and Thin Gap Chambers (TGC) in the end-cap region. In addition, charged particle tracks are bent by the barrel and end-cap toroid magnets described in the next section.

The LHC bunch spacing is 25 ns. This is smaller than a typical muon time-of-flight across the ATLAS detector which is around 70 ns. The trigger chambers have good

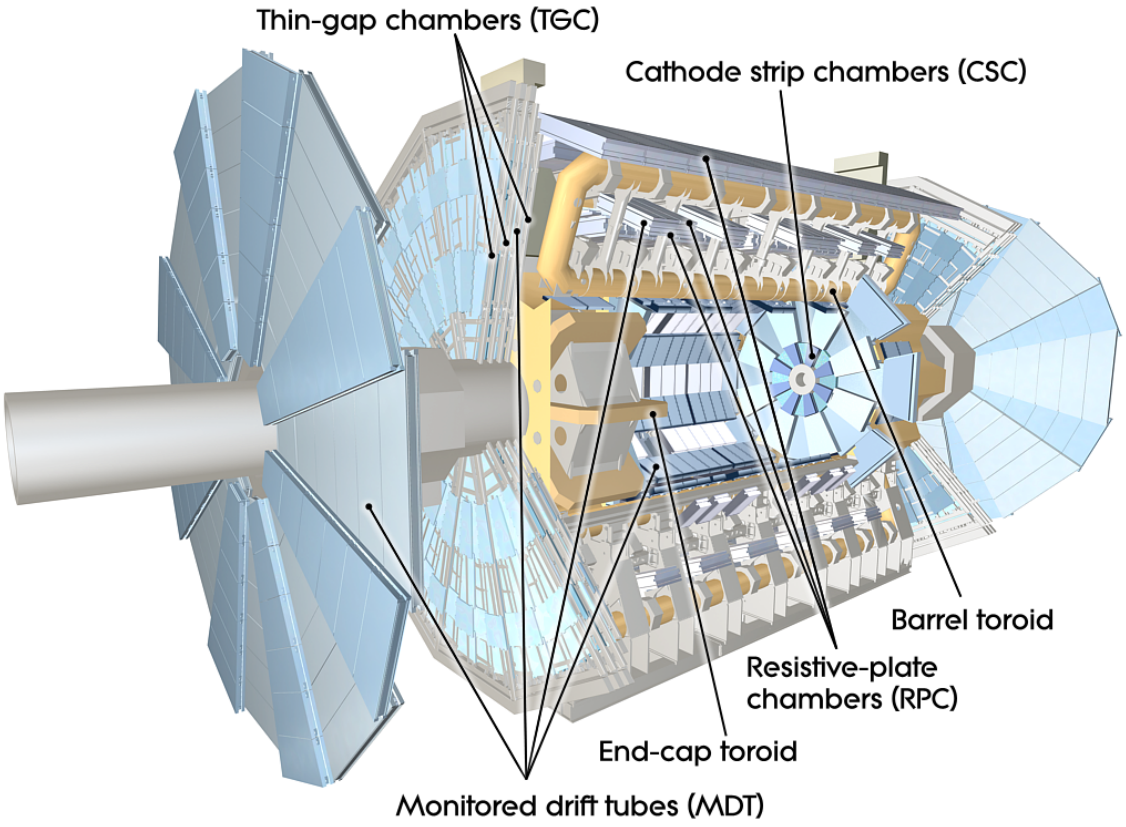


Figure 3.7: Schematic drawing of the ATLAS muon system [48].

timing resolution which enables them to link a particular muon track to a specific bunch crossing. The precision chambers provide good spatial resolution which is used together with the timing information from the trigger chambers to create muon tracks in a specific event. Table 3.1 provides a summary of the covering spatial and timing resolutions of the various tracking and trigger chambers. For high p_T muons with low track curvature, the trigger chambers provide some coarse spatial information for use in calculation of the track parameters. Table 3.1 shows the typical number of muon hits in the barrel and end-cap regions, with some extra hits coming from the trigger chambers. The majority of tracking chambers are MDTs and these are spread in three barrel layers (stations) and three outer end-caps (the inner most end-caps are comprised of CSCs). The MDTs are aluminum tubes of diameter 3 cm containing a central axis wire which works as an anode. The tubes are filled with 97% Ar gas and 3% CO_2 at high pressure. A typical muon passing through barrel or end-caps will give 20 hits in the MDTs.

The CSCs are radiation hard multi-wire proportional chambers and provide (on average) 4 hits per muon in the first wheel of the MS.

			Chamber resolution (RMS)			hits/muon	
Type	**Function**	**Coverage**	z/R	ϕ	time	barrel	end-cap
MDT	tracking	$\|\eta\| < 2.7$	35 μm (z)	—	—	20	20
CSC	tracking	$2.0 < \|\eta\| < 2.7$	40 μm (R)	5 mm	7 ns	—	4
RPC	trigger	$\|\eta\| < 1.05$	10 mm (z)	10 mm	1.5 ns	6	—
TGC	trigger	$1.05 < \|\eta\| < 2.7$	2-6 mm (R)	3-7 mm	4 ns	—	9

Table 3.1: The different parameters of the muon subdetectors. Spatial (columns 4 and 5) and timing (column 6) resolution are presented. The last two columns show the number of hits per track [48].

3.2.4 Magnet System

The magnet system with its two parts, solenoid and toroid magnets is considered to be one of the principal components of ATLAS and the one which sets ATLAS apart from other detectors. The magnet system is comprised of a central superconducting solenoid surrounding the ID, and three groups of air-core toroids, also superconducting. The toroid magnet system is composed of a barrel section and two end-caps. The barrel and end-caps are each made up of eight individual rectangular coils arranged as shown in Figure 3.8. The magnetic fields are mostly orthogonal to the outgoing particles which allows ATLAS to obtain a precise measurement of the charge sign and momentum.

The solenoid magnet bends the charged particles in the $x - y$ plane by providing a magnetic field of up to 2.6 T causing low energy particles to spiral inside the ID where one observes tracks in the pixel, SCT and TRT layers. The high energy particles have a straight trajectory and they escape from the solenoidal field to the outer layers of ATLAS such as the calorimeters and MS. The toroid magnet system provides a varying magnetic field with strengths up to 4.1 T in the region of the MS by using a system of 8 rectangular coils in the barrel and end-caps constructed radially. Each part of the toroidal magnet bends the charged particles in the $\pm z$ direction. To enable the superconductivity, solenoidal and toroidal magnet systems are cooled with liquid helium to a temperature of 4.5 K.

3.3 Trigger and Data Acquisition

The LHC provides bunch crossings to ATLAS every 25 ns (40 MHz). These bunch crossings produce an interaction rate of about 1 GHz due to the pileup activities.

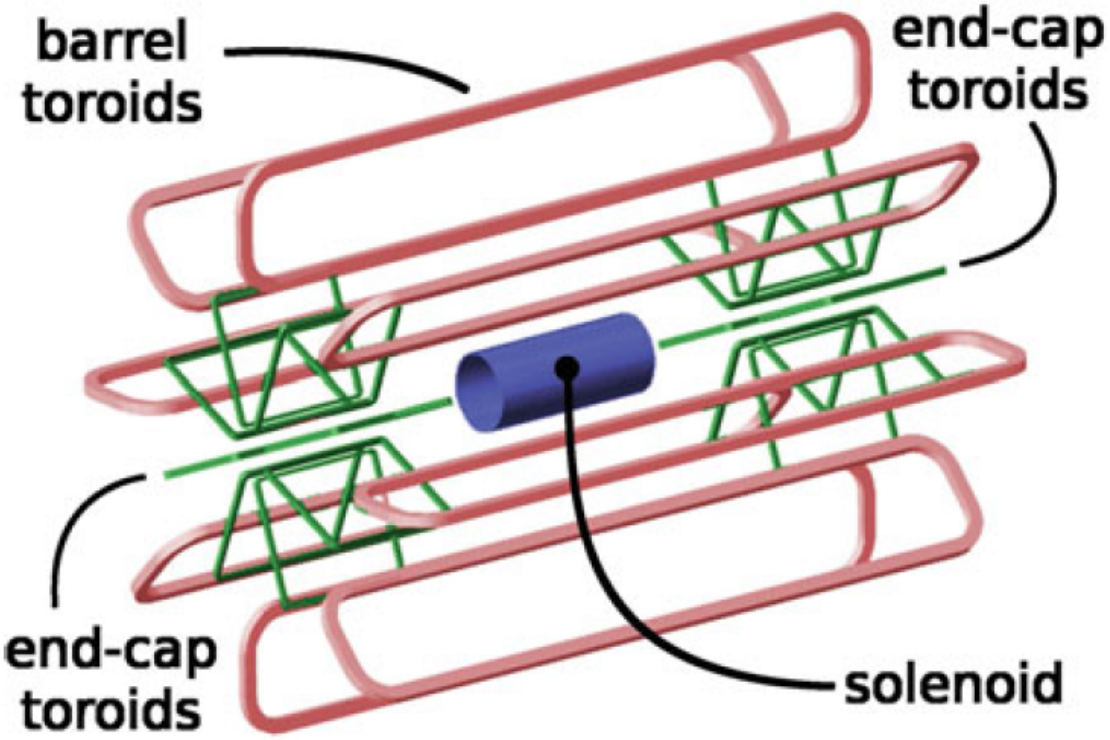


Figure 3.8: Schematic drawing of of the ATLAS magnet system: three external toroids comprising in MS and central solenoid covering ID [48].

The mean data size readout in ATLAS for each bunch crossing is approximately 1-2 Mb. The storage cost of this huge amount of data would be very high, however, the vast majority of events do not include interesting physics processes. Ideally, all the interesting data needs to be recorded. To attain this, ATLAS has a three level trigger system to filter interesting events as shown in Figure 3.9. These three levels consist of a hardware-based level trigger *Level-1* (L1) and a software-based *high level trigger* (HLT) which includes *Level-2* trigger (L2) and the *Event Filter* (EF).

The L1 trigger receives coarse information (with coarse resolution) from the calorimeters and MS trigger chambers (RPCs and TGCs). At this level, the trigger reduces the event rate from 40 MHz to about 75 kHz. This level uses conventional electronics to select the interesting (high p_T) bunch crossings. When the L1 trigger is passed, it also determines interesting features for the candidate event which form the *Regions-of-Interests* (RoIs). The event then passes to the next trigger stage HLT. At the L2 stage, the candidate jets, hadrons, high p_T electrons and muons and photons are calculated. Given this, all of the data have to processed as fast as possible. Therefore, the decision is targeted to take at most 2.5 μs and typically 100-120 bunch crossing are kept in the pipeline memory.

The L2 trigger is a computing farm system which takes into account the full-

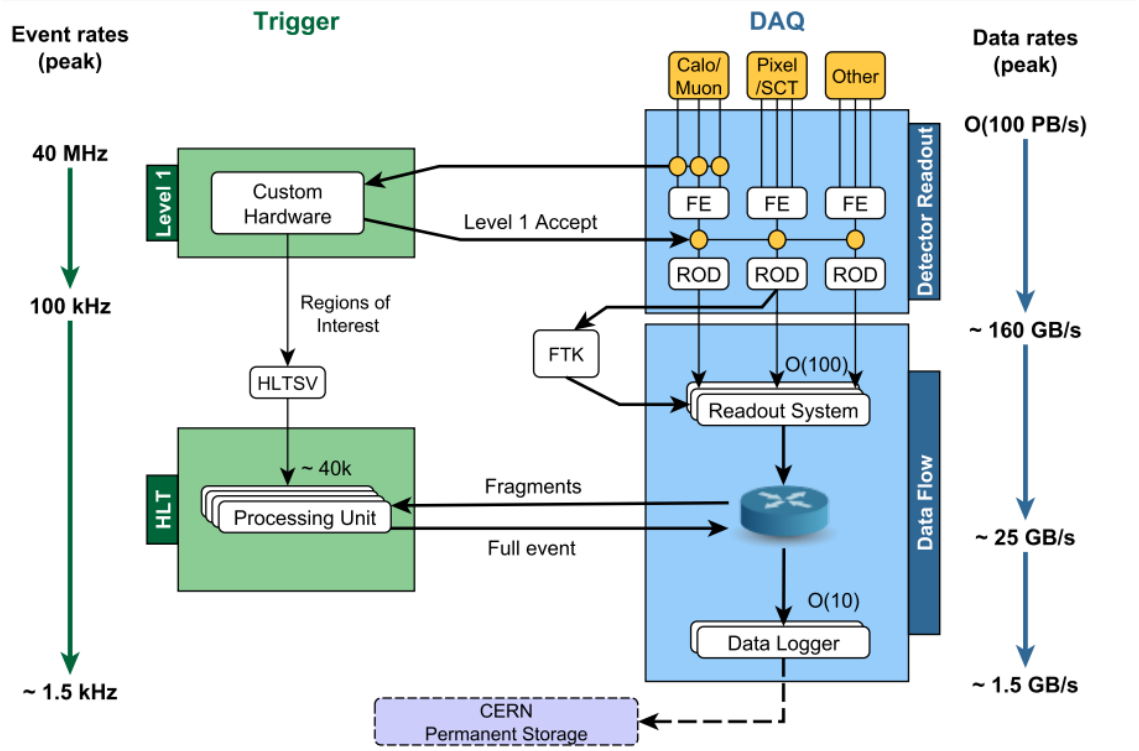


Figure 3.9: Schematic drawing of of the ATLAS trigger system [59].

granularity data and to the information from the ID. To preserve the decision time performance which its average processing time per event is 10 ms, the L2 trigger only analyzes the RoI's using the dedicated trigger algorithms. The reduction of the event rate by the L2 trigger is to the order of 1 kHz.

Finally, events passed to the third and final trigger level EF which are fully reconstructed using ATLAS offline event reconstruction software and thus, the EF has no need to use dedicated reconstruction algorithms. This last decision has a latency time of the order of a second. The EF reduces the data rate to about 100 Hz and the events are stored for later use in physics analysis.

Chapter 4

Analysis strategy and techniques

The search for the $t\bar{t}H(H \rightarrow bb)$ process faces several difficulties due to the large backgrounds, low cross-section compared with pp→VH or pp→H channels, and its final state being complicated by the presence of 6 non-overlapping high-p_T jets. The jets include 4 b-jets and 2 light quark jets: this gives rise to a challenging analysis in two ways. The high combinatorics from the 4 b-jets and the large backgrounds from $t\bar{t}$+light jets. The result is that the small signal to background ratio after the event selection and categorization as signal-rich and background-enriched regions that are explained in section 4.1. In order to increase the purity of the signal-rich regions; $(\geq 6j, \geq 4b)$, $(\geq 6j, 3b)$ and $(5j, \geq 4b)$, it is essential to optimize the sensitivity of the search by using a powerful discriminating variables through the application of a *multivariate analysis* (MVA) which is described in Section 4.2. These signal-rich regions have been analysed using a likelihood fit to evaluate the signal strength μ in section 4.3.

4.1 Analysis strategy and event categorisation

This thesis presents a search for the SM Higgs boson produced in a association with top quarks decaying to b-quark pair where one of the top quarks decays leptonically (e or μ with corresponding ν) and the other top decays hadronically, the $t\bar{t}H \rightarrow (l\nu b)(jjb)(b\bar{b})$ signal lepton channel is considered. Several advanced approaches are adopted to extract and identify this signal because of the low $t\bar{t}H$ production with respect to a wide variety of background processes. In order to optimize the sensitivity of the search, the selection events are categorized into different regions. Depending on the number of jets with p_T >25GeV and the number of b-tagged jets giving

region with m jets and n b-tagged jets (mj,nb). Nine independent topology regions are considered in this search. The regions with a large expected "signal" (S) to "background" (B) ratio S/B >0.5% and $S/\sqrt{B}$ >0.5 are referred to as "signal-rich" regions. These regions are, ($\geq$ 6j,$\geq$ 4b), ($\geq$ 6j,3b) and (5j,$\geq$ 4b) which provide most of sensitivity to the signal. The remaining six regions are referred to as "signal-depleted" or "control" regions. They are very pure background-only regions and are used to constrain systematic uncertainties to improve the background prediction in the signal-rich regions. These regions are, (4 jets, 2 b-tags), (4 jets, 3 b-tags), (4j,4b), (5j,2b), (5j,3b), ($\geq$6j,2b). Figure (4.1) shows the expected S/B ratio for the different regions of the standard analysis variables of semileptonic $t\bar{t}H$ channel. The background composition across different jets and b-tags multiplicity regions are illustrated in Figure (4.2).

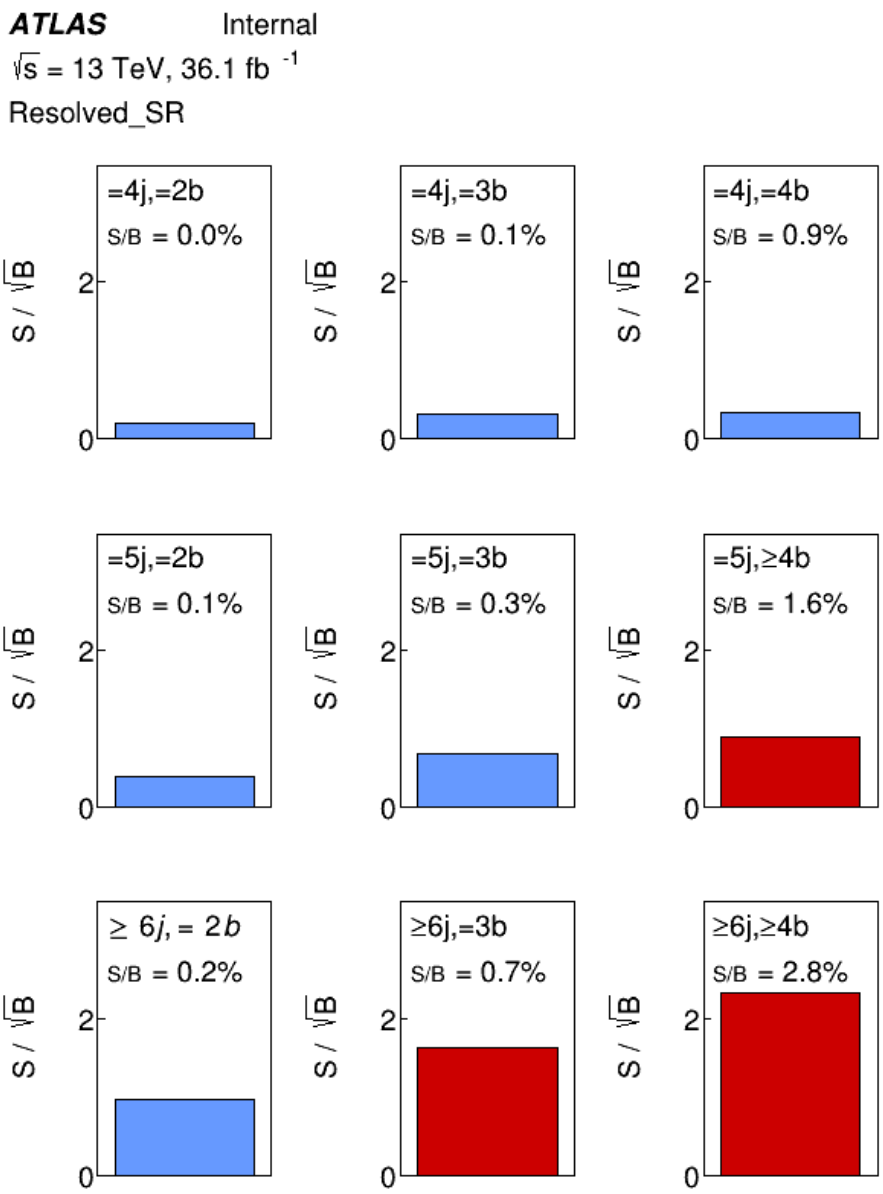


Figure 4.1: The S/B and S/$\sqrt{B}$ ratio for each of the regions assuming SM cross-section and m_H=125 GeV. The rows show the plots for the jet multiplicity (4, 5, ≥6), and the columns show the b-jet multiplicity (2, 3, ≥4). The dark red illustrates signal-rich regions while the light blue refers to the signal-enrich regions.

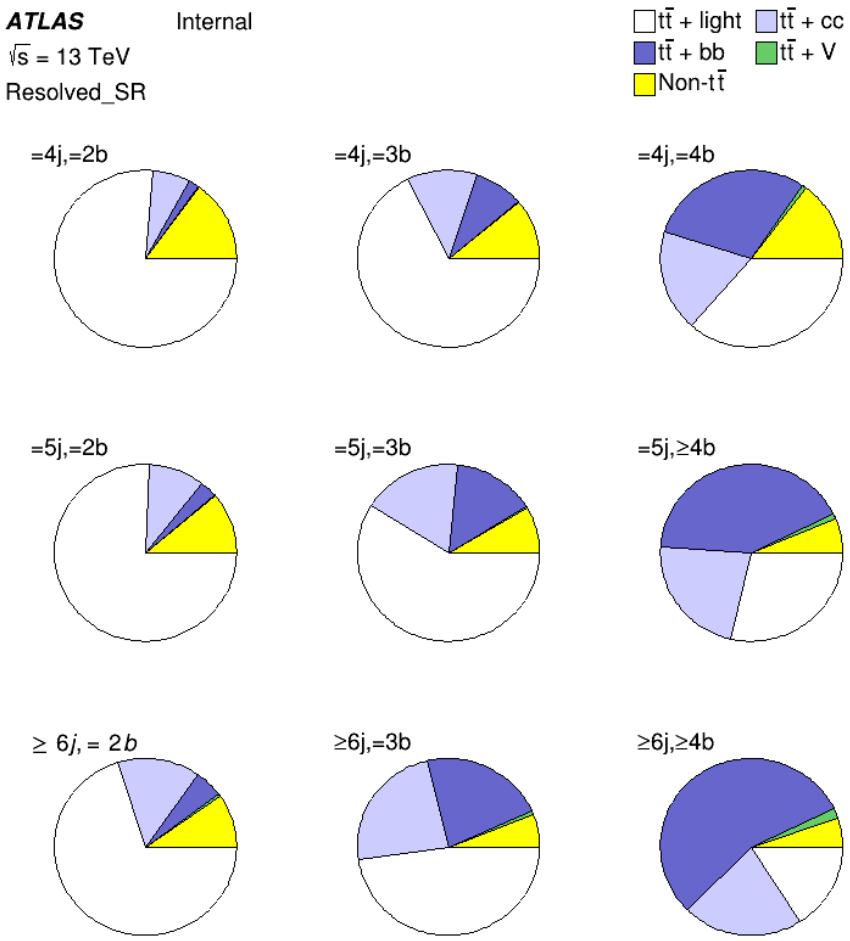


Figure 4.2: Background composition for the resolved analysis regions after pre-selection cut.

In the most signal rich region (≥6j,≥4b), the signal significance reaches $S/\sqrt{B}$ ≈2 and the signal to background ratio is 2.8% which means the cut-and-count approach to extract the signal is a bad solution and leads to a tiny amount of signal and large systematic uncertainties on background estimation. The idea that is followed in this thesis is to implement a kinematic fit of a given event topology to a kinematic model using a likelihood approach, KLFitter, in the most sensitive region (≥6j,≥4b). The maximisation of the likelihood is done using the the BAT package [60] as is explained in more detail in the next chapter. The kinematic fit then tries to reconstruct Higgs (H→bb) from the measured objects. In doing this, it provides some output variables that are directly sensitive to how signal-like the selected events are. After that, the TMVA (for more detail, see Section 4.2) takes the output from the KLFitter as input variables to discriminate signal from background in the region (≥6j, ≥4b). For the final measurement, these variables can be combined with the non-KLFitter variables

illustrated in table (9.8) (from ICHEP [61]) which are used in the corresponding region and extracted from TMVA which is detailed in section 9.1. In the end, this KLFitter region is combined with the non-KLFitter regions (the remaining eight regions) to set the upper limit $\mu = \frac{\sigma_{observed}}{\sigma_{expected}}$ (see 4.3) by fitting the shape of the MVA output trained using TTHFitter package, which is used in the ATLAS $t\bar{t}H$ group to perform the fit to data, see section 9.5.

4.2 MultiVariate Analysis technique

4.2.1 Machine Learning Technique

An MVA combines the information from specific input variables which exhibit a clear separation between signal and background into one output discriminant variable. In this case, the output is a parameter indicating that an event is signal or background. The input variables in the MVA will be those likely to be able to differentiate S/B. Choice of the variables for the MVA is done through "ranking" procedure. Ranking gives a numerical value for each variable indicating how well that variable distinguish S/B. The variables that are used in the MVA have one value per event and then, the values for each event for each variable are summed. In the signal rich regions, the $t\bar{t}$+bb is the dominant background and in the other signal-depleted regions, $t\bar{t}$+light jets dominates. Obviously, any variable which allows differentiation of $t\bar{t}$+bb from $t\bar{t}H(H \to bb)$ in the signal-rich regions is vital. For example, if the bb from $t\bar{t} + bb$ differ in some way from the $H \to bb$. The MVA is not trained on the control regions, they are fitted using H_T^{had} variable which is defined in table (9.8). The region (5j,3b) is the only exception to the control regions that the MVA is trained on this region to discriminate $t\bar{t}$+HF jets and $t\bar{t}$+light jets by using the third b-tag variable.

To discriminate the correct jet-parton assignment from the combinatorial background, a supervised *Boosted Decision Trees* (BDTs) technique are used which is implemented in the TMVA package [11].

Boosted decision trees

BDTs are an example of a multivariate algorithm used for particle and event classification. It is a relatively recent development in machine learning which has in recent years become widely used in the LHC. This is because it is easier to tune and the input variables do not require adaptation when they are used within BDTs. A

decision tree is a binary tree structured classifier where repeated yes/no decisions are taken on an MC sample as shown in Figure 4.3.

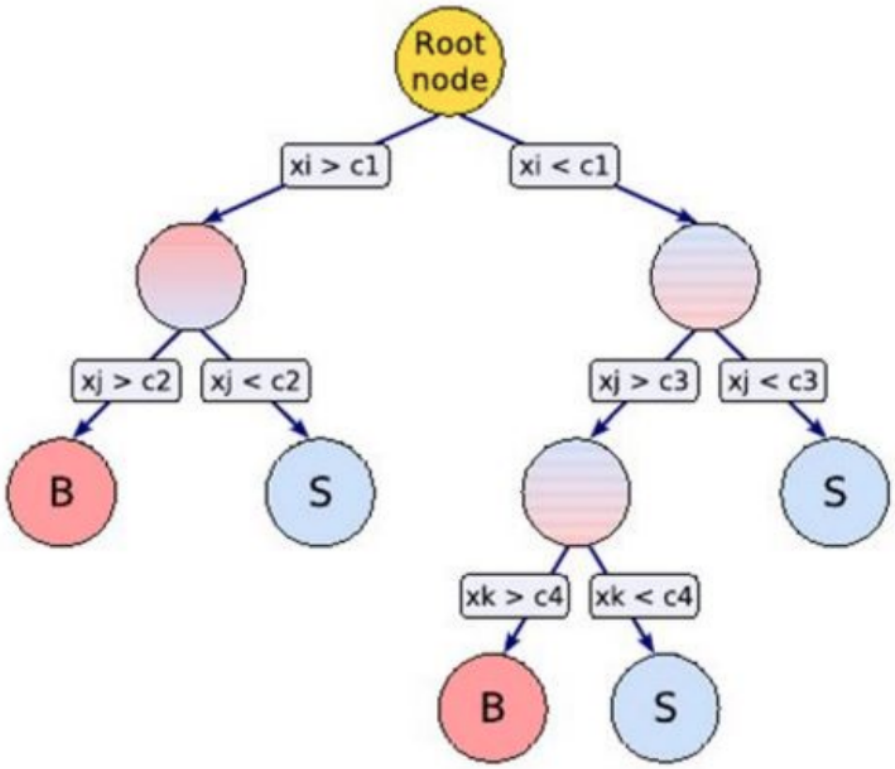


Figure 4.3: Schematic drawing of a decision tree. Starting from the root node, a sequence of binary splits on events starting with the discriminating variable x_i.

The MC signal and composite background events are divided into two samples (sample A and sample B) for training and testing based on the event number. Therefore, a signal decision tree is trained on one sample and tested on the other sample. The root node uses only a single discriminating variable x_i which specifies the best cut value that leads to best S/B separation. In other words, this variable creates two new nodes (two new sub-samples). One of them is signal like and the other is background like. The mechanism of discrimination that is described for the root node is repeated on the two new nodes but with the using of different variable x_j. This new variable specifies two different "cut criteria" for each node which split it further into two samples. This process continues until a stop criterion is fulfilled based on the minimum or maximum S/B ratio or number of events in a node.

BDTs are susceptible to overtraining that the decisions with a decision tree are based on the best cut value at each stage which is influence by the statistical fluctuations of variables values in the training population. If the training population is too small and the number of variables are too high, the tree may be overtrained (events are

memorized).

"Boosting" represents a practical solution to improve the performance and stability of decision trees. In particular, overtraining is reduced by boosting. The type of boosting chosen for this analysis is the *Adaptive Boost* (AdaBoost) which is described in the following lines. After the first tree is trained, the misclassified events (e.g the background events from training end up in signal and vise virus) are giving a higher weight. A second tree is then trained on the re-weighted sample. The second tree focuses on the events which have been classified wrongly in the previous tree. In the second tree, the weights of wrongly classified events are multiplied by a common boost weight factor α eq.(4.2) which is derived from the misclassification rate, err

$$\text{err} = \frac{\text{misclassification events}}{\text{total events}}, \tag{4.1}$$

$$\alpha = \frac{1 - \text{err}}{\text{err}}. \tag{4.2}$$

This process is repeated many times (typically 100-500 trees). To verify that there is no overtraining, the performance of training and testing samples should be similar. To conclude the idea of boosting; n bootstrap copy data sets are made from the original training sample and each trains an independent Decision Tree. These bootstrap samples appear as different samples (because of the weighting) to the original one which means it is hard for the tree to memorize the events. Finally, the BDT is obtained by combining the responses of the set of DTs.

4.2.2 Kinematic and Event Topology variables

In the analysis that is used in this study, the training has been performed with a set of variables that are detailed in tables 9.1 and 9.8. BDTs are used to discriminate signal from the background in the single lepton channel. Also, these variables are used as a base to study the separation performance of KLFitter variables which is explained in detail in the next chapter. Several classes of variables have been inspected for their discriminating power considering the variety of regions and the rich topology of the events. These variables are classified into the following classes:

- Object kinematic variables: including p_T and η of the lepton, and of each of the jets such as p_T^{jet5}, see Table 9.1.

- Global event variables: they summarize for each event information on kinematics such as the scalar sum of the jets p_T's (H_T^{had}), see Table 9.8.

- Event shape variables: they are complex event variables that make use of the full final state which are the combination of the eigenvalues of the linear momentum tensor [62] like $H1$, Aplanarity and Centrality, see Table 9.1.

- Object pair properties: including combinations of two objects with smallest or largest values of ΔR, p_T vector sum and invariant mass. This class comprises the rest of the 14 variables used in the BDT analysis.

4.2.3 Classification performance evaluation

There are several classification techniques and settings which can be used to assess the performance and the discriminatory power of the BDTs. These different benchmark quantities have to be evaluated independently and a brief explanation of each of them is mentioned in the following;

- The *separation* $< S^2 >$ of a classifier y which is defined as:

$$< S^2 >= \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy \tag{4.3}$$

  where $\hat{y}_S(y)$ and $\hat{y}_B(y)$ are the signal and background probability density function of y, respectively.

- The discrimination *significance* of a classifier which is defined by the difference between the classifier means for signal and background divided by the quadratic sum of their root-mean-squares.

- BDTs setup: to get the best significance and separation power, as an example of the optimal set of BDTs parameters are chosen as shown in table (9.2).

- The distribution of signal and background shapes of the input variables shows the discriminating power of each of them. An example is the centrality variable which has the highest discrimination significance and separation power, as illustrated in Figure 4.4.

- The correlation between two random variables X and Y is measured with the correlation coefficient ρ which is defined as:

$$\rho(X|Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{4.4}$$

where $cov(X, Y)$ is the covariance between the two variables, and $\sigma_X^2(\sigma_Y^2)$ is the variance of the variable X(Y). An example of the correlation among the variables depending on eq.(4.4) for signal and background, are illustrated in Figure 4.5.

- The separation power and the importance ranking are the factors taken into account in order to chose the final set of the discriminating variables.

- Kolmogrov Smirnov (KS) test: the BDT output distribution for testing and training which is also called BDT response, see Figure 4.6, can be used as input to the KS test as an overtraining check. This test provides a p-value of the training and testing samples drawn from the same distribution. The BDT output distribution are considered to be overtrained if its KS value is below the critical value KS$\lesssim$0.01[63].

- Cross-training test : another efficient way to check the overtraining is the cross-training test. The training signal and background samples (A and B) are divided into two equal sub-samples based on the event number that sample A takes the even event numbers and sample B takes the odd event numbers. Then, one is used as a testing sample, while training is performed on the other sub-sample, respectively. The corresponding ROC curves of the two classification plots obtained are overlaid and compared. If the ROC's shapes appears to be equivalent for the two evaluations, the test is assumed to be passed. An example is shown in Figure 4.7.

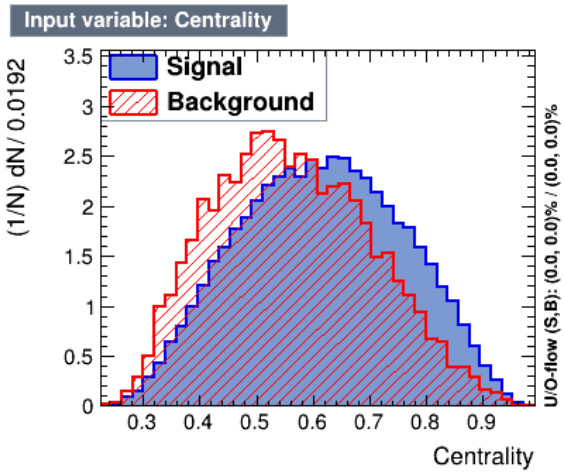
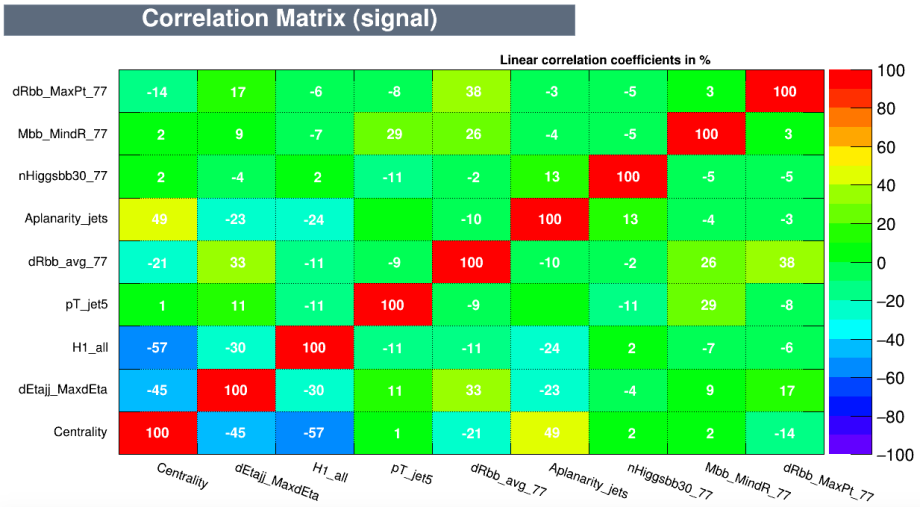
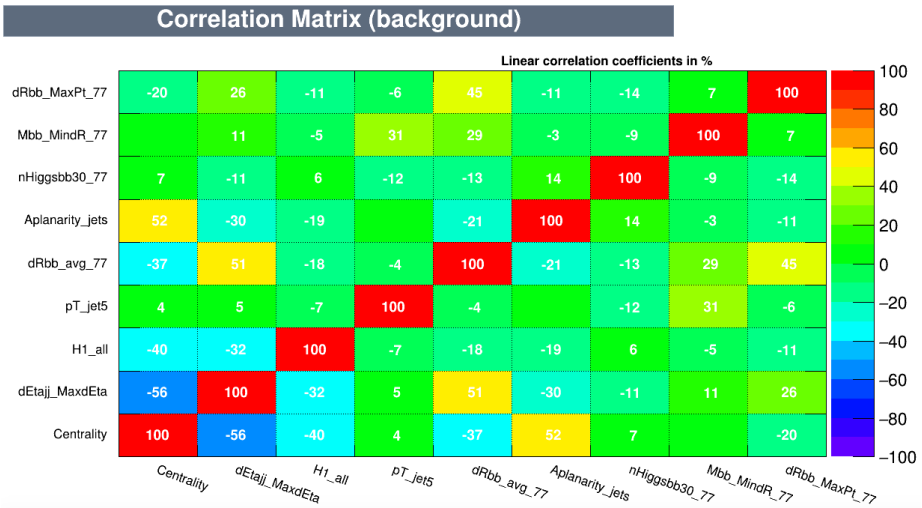


Figure 4.4: The distribution of the input centrality variable for the training sample in the region ($\geq$6 j, $\geq$4 b).



(a)



(b)

Figure 4.5: Correlations between input variables for the BDT training for signal (a) and background (b) in the region ($\geq$6 j, $\geq$4 b).

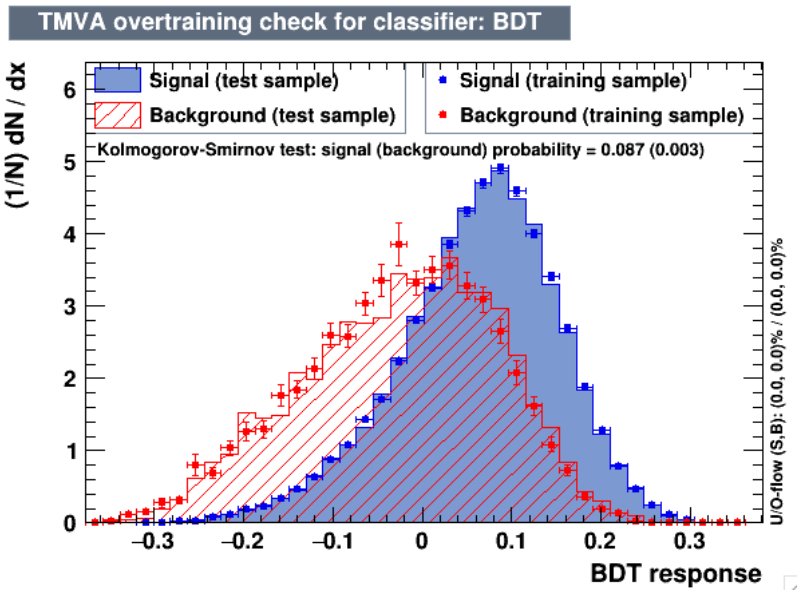


Figure 4.6: BDT responses for signal (red) and background (blue) samples in the region (≥6 j, ≥4 b).

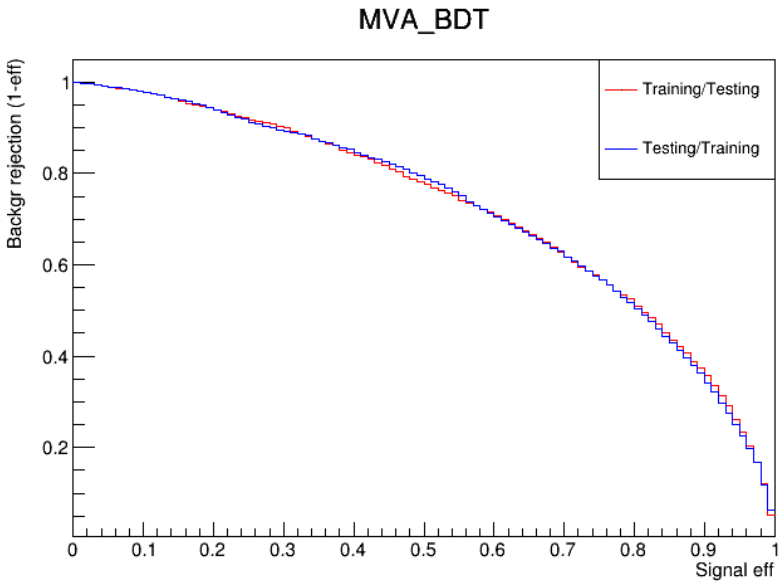


Figure 4.7: ROC Curves for the cross-training test in the region (≥6j,≥4b).

4.3 Statistical analysis

Statistical tests are an important tool to determine weather the discovery or exclusion of a new physics model in particle physics is compatible with SM prediction or not. The testing for presence or absence of signal, assuming a Higgs boson mass of m_H=125 GeV, is achieved by comparing the consistency of the data with background-only hypothesis H_0 (null hypothesis), or signal-plus-background hypothesis H_1 (test hypothesis) with the inclusion of the Higgs boson as a signal. The two hypotheses can be generalized by introducing the parameter of interest μ (signal strength factor) which is defined as the ratio of the observed $t\bar{t}H$ cross-section to the $t\bar{t}H$ SM expected cross-section:

$$\mu = \frac{\sigma^{\mathrm{ttH}}_{\mathrm{observed}}}{\sigma^{\mathrm{SM\ ttH}}_{\mathrm{expected}}} \tag{4.5}$$

The factor μ quantifies the multiple of the signal cross-section "μ signal cross-section + background-only" (μs+b). In other words, the expected number of events E equal μs+b. The background-only hypothesis H_0 assumes the absence of $t\bar{t}H$ signal and therefore μ=0. Conversely, the "test" hypothesis ($H_1 = 1\times$s+b) corresponds to μ = 1. By assuming the data that is used in this thesis follows a Poisson distribution with parameter λ eq.(4.6), the corresponding expectation value of the Poisson distribution is also λ

$$P(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \tag{4.6}$$

where $P(x)$ is the probability of observing x events in a given interval. By describing the number of expected events in each bin E_i

$$E_i = \mu s_i + b_i, \tag{4.7}$$

and the number of observed events in a bin i as n_i, the likelihood for the observed data to be produced by the model in the i^{th} bin is described by the following equation

$$L_i(\mu) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}, \tag{4.8}$$

where s_i and b_i represent the number of expected signal and background events in each bin i.

An overall estimate of how likely this data will be produced by the model is calculated using a likelihood function $\mathcal{L}$ which is the product of Poisson probability terms over all bins N

$$\mathcal{L}(\mu) = \prod_{i=1}^{N} \frac{(E_i)^{n_i}}{n_i!} e^{-E_i}. \tag{4.9}$$

The model includes uncertainties because it is a stochastic system. These uncertainties in the form of systematic and statistical errors can be modeled through *Nuisance Parameters* (NP), θ_k. The NP are implemented in the likelihood function as a different *Probability Density Functions* (PDF's). In this analysis, three PDF forms are used, corresponding to two kinds of systematic uncertainties, shape and normalization, and statistical uncertainties which together encode the effect of uncertainties on s_i and b_i expectations. In the limit of small uncertainties, all the PDFs approximate to a Gaussian distribution centered around zero with a width (σ) of one (standard normal distribution) in order to simplify the analysis. In general, systematic uncertainties are assumed to have a Gaussian prior $\rho(\theta_k)$. For M Nuisance Parameters, the full likelihood can be written as:

$$\mathcal{L}(\mu) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \prod_{k=1}^{M} \rho(\theta_k). \tag{4.10}$$

Therefore, the values of μ and θ_k give the total number of expected events in a given bin. Moreover, the changes to the systematic uncertainties (NP's) made by the fit adjust the expectations for signal and background to give the best fit to data. The best estimate for μ and θ is obtained by maximizing the likelihood. In fact, the ratio of *two* likelihoods is used. One of the likelihoods is globally maximized with no constraint on μ or θ, and is defined as $\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})$, and the other likelihood $\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)$ is maximized for a fixed value of μ. Together these form a profile likelihood ratio

$$\lambda(\mu) = \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \tag{4.11}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximize the unconditional likelihood function with the constraints $0 \leq \hat{\mu} \leq \mu$. The conditional likelihood function $\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)$ can also be maximized for a fixed value of μ giving the estimate of the NP as a function of μ, $\hat{\theta}_\mu$.

The profile likelihood ratio is then is used in the calculation of the *test statistic* q_μ:

$$q_\mu = -2\ln\frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta})}, \text{ with a constraint } 0 \leq \hat{\mu} \leq \mu \tag{4.12}$$

The advantage of using the profile likelihood ratio is to simplify the analysis because for a large sample, the distribution of $q_\mu = -2\ln\lambda(\mu)$ has a chi-square shape.

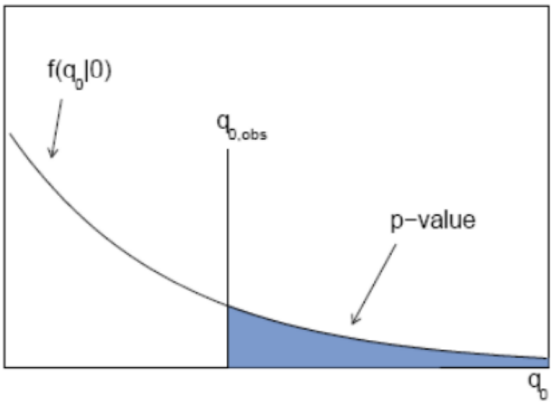
To quantify the level of disagreement between the data and the hypotheses of μ=0,1, the p-value is computed

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu^{'})\mathrm{d}\mu, \;\; \mu \neq \mu^{'}. \tag{4.13}$$

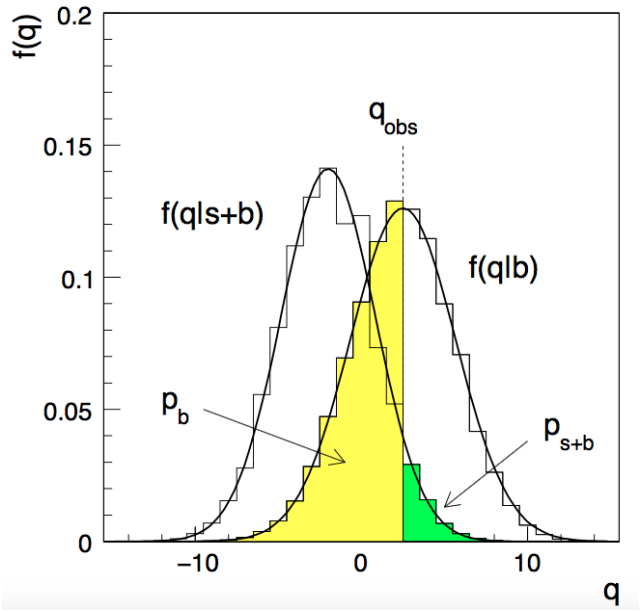
It is defined as the probability that the observed data $q_{\mu,obs}$ originates from the considered hypothesis.

Therefore, the significance for a discovery ($f(q_0|1)$) is computed as a number of standard deviations Z corresponding to an upper-tail probability of p_μ for a Gaussian distributed variable of zero mean as shown in Figure 4.8a. The discovery is claimed if the p-value of the background-only hypothesis is below 2.9×10^{-7} corresponding to a significance Z=5. However, the expected 95% CL for the exclusion ($f(q_1|0)$) of signal-plus-background hypothesis is computed by using the following probability ratio

$$CL_s(\mu) = \frac{CL_{s+b}}{CL_b} = \frac{p_1}{1-p_0}. \tag{4.14}$$



(a)



(b)

Figure 4.8: (a) Illustration of the the p-value of discovery corresponding to the median of q_μ assuming a strength parameter $\mu^{'}$. (b) The distribution of the statistic $q_\mu = -2\ln(\mathcal{L}_{s+b}/\mathcal{L}_b)$ for b-only and s+b hypotheses [64].

where p_1 and p_0 quantify the agreement between the data and signal-plus-background and background-only hypothesis respectively, as illustrated in Figure 4.8b. The 95% upper limit on μ is achieved by adjusting μ until the value of CL_s=0.05 is reached. The most recent central fit value of the signal strength parameter for m_H=125 GeV is $\hat{\mu} = 2.00 + 1.09/$ -0.95. As for the observed data, the value of the signal strength parameter μ which gives CL_s=0.05 is 3.9 [8].

Chapter 5

Object reconstruction

The final states of interest in this thesis contain a hard and isolated lepton (electron and muon), a neutrino giving rise to missing transverse momentum, two light quarks jets and four b-quark jets. This chapter describes the reconstruction of these objects from the raw detector data.

5.1 Tracks and Vertices

Tracks produced by the charged particles traversing the ATLAS detector along a circular trajectory in the transverse plane of the Inner Detector (ID) solenoidal magnetic field. They are reconstructed within the full ID acceptance range ($|\eta|$ <2.5) and their charges and momenta are measured. The accurate reconstruction of tracks that results from the charged particles trajectories mandatorily require good pattern recognition algorithms such as a global χ^2 fitter [65]. The pattern recognition works primarily inside-out. It is designed to reconstruct the primary charged particles. It starts building track "seeds" considering space points in the silicon detectors and then extrapolates the trajectory to the TRT. The outside-in algorithm is used to reconstruct tracks from secondary interactions which represent the signals in the TRT that are not associated to any track candidate by the inside-out reconstruction. This algorithm uses the TRT track segments as seeds and a track candidate then is extended to the silicon detectors.

The ID is also designed to reconstruct the primary and secondary vertices using an iterative vertex finding algorithm [66]. The reconstruction of the *Secondary Vertices* (SV) and displaced tracks is used to tag the jets originating from the long-lifetime particles such as c-quark and b-quarks. The number of reconstructed vertices is a

direct measure of the amount of in-time pile-up [67]. The track reconstruction has been optimised in Run2. The seed purity is improved and the ambiguity-solving method in dense environments is updated [68].

A reconstructed track is fully specified by a set of parameters with respect to the primary vertex: the transverse (d_0) and the longitudinal (z_0) impact parameters, the azimuthal (ϕ) and polar (θ) angles and the inverse transverse momentum (q/p_T) where q is the particle charge, see Figure (5.1).

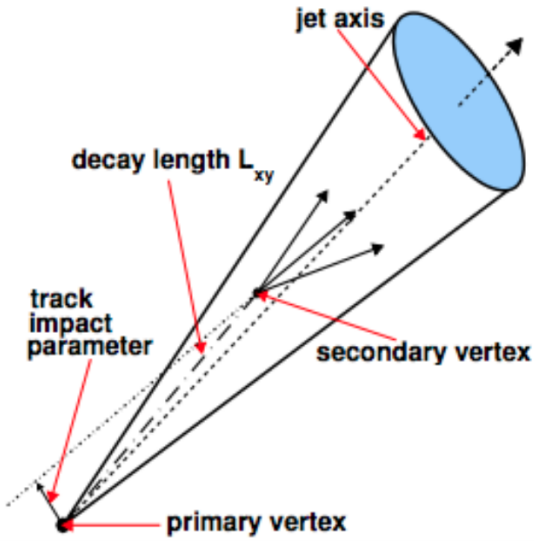


Figure 5.1: Illustration of the geometric set of track parameters [69].

5.2 Electrons

Electron-finding in ATLAS occurs in two steps; *electron reconstruction* and *electron identification*. Electron reconstruction provides a set of electron candidates with high transverse momentum $p_T > 27$ GeV. These electrons are seeded by clusters of deposited energy in the ECAL which are associated with a reconstructed track in the ID. Reconstruction is performing using the sliding window algorithm which is only used for electron and photon reconstruction in ATLAS [70]. This algorithm performs a scan of the calorimeter by searching for a local maximum of energy within a rectangular window of fixed size. The cells within this window then form the cluster. The full four-momentum is then constructed using track η and ϕ and the cluster energy.

The electron candidates have in addition to real electrons, "fake" electrons which arise mostly from hadrons that are misidentified as electrons (multijet). Therefore,

to identify the real electrons out of these candidates and reject as many of the fake electrons as possible, electron identification is required. The baseline electron identification in ATLAS uses cuts on several discriminating variables. three reference sets of cuts are defined for use in analysis, denoted "loose", "medium" and "tight" in order to increase the background rejection. The identification algorithms are based on a multivariate likelihood discriminator which uses signal and background probability density functions (PDFs) of the electron discriminating variables [70]. The analysis presented in this dissertation use the tight likelihood trigger electron identification (TightLH) since the largest possible "fake" electrons from mis-identifications will be rejected. Electrons are required to have $p_T > 10$ GeV and $|\eta_{\text{cluster}}| < 2.47$ and the calorimetry transition region $1.37 < |\eta_{\text{cluster}}| < 1.52$ is excluded since this region shows worse energy resolution performance and reconstruction. Further selections on the transverse and longitudinal impact parameters are made: $|d_0/\sigma_{d_0}| <5$ (σ_{d_0} is the uncertainty on the reconstructed d_0) and $|z_0 sin\theta| <0.5$ mm.

Isolation criteria are not used in the electron identification. However, isolation criteria are an important tool in order to isolate electrons that are produced at the primary vertex. In this case, the other sources of real electrons like those from semi-leptonic decays of heavy hadrons and photon conversion are not isolated and thus are not used in the reconstruction of the $W^{\pm}$. Two main kinds of variables are designed for this discriminating purpose; the *calorimetric isolation energy* $E_T^{cone0.2}$ and the track isolation, $p_T^{varcone0.2}$ [70]. The $E_T^{cone0.2}$ defined as the sum of transverse energies in a cone of ΔR=0.2 around the candidate electron cluster, excluding the electron energy itself. The $p_T^{varcone0.2}$ defined as the sum of transverse momenta of all tracks in a cone of radius ΔR=0.2 around the candidate electron track and originating from the reconstructed primary vertex of the hard collision, excluding the electron track and additional tracks from converted bremsstrahlung photons. Electron candidates to be isolated are also required to pass the "Gradient isolation working point [71] which gives several isolation efficiency values by using different p_T values. Electron candidates must pass the Gradient isolation working point corresponding to efficiency $\epsilon = 0.1143 p_T[\text{GeV}] + 92.14\%$ (equal to 90% for $p_T = 25$ GeV and 99% for p_T =60 GeV, respectively. The efficiency plateaus out at 100% for electrons having p_T >60 GeV).

5.3 Muons

As mentioned in Section 3.2.3, the muon spectrometer (MS) is designed to detect muons in the pseudorapidity region $|\eta| < 2.7$. In ATLAS, several types of reconstructed track in the MS+ID are available [72]. The analysis in this dissertation uses only the combined (CB) muon. In more detail, the muon reconstruction is performed separately in the inner detector (ID) and MS. A combined muon is created by combining the reconstructed track in the MS with a matching ID track. A combined track is formed by a global refit of the hits of both tracks, taking into account the muon energy loss in the calorimeter. The reconstruction of muons mostly follows an outside-in pattern recognition where the muons are first reconstructed in the MS and then extrapolated towards to match an ID track.

To identify the muon that is expected as *signal* muon if it results from the decay products of W boson, it must meets specific conditions and cut criteria applied to the variables used in muon identification. These conditions and variables are:

- q/p *significance*, defined as the absolute value of the difference between the ratio of the charge and momentum of the muons measured in the ID and MS divided by the sum in quadrature of the corresponding uncertainties.
- $\rho^{'}$, defined as the absolute value of the difference between the transverse momentum measurements in the ID and MS divided by the p_T of the combined track.
- normalized χ^2, using to fit the hits associated with each track .
- The number of hits in the ID and MS must be meet the following requirements:
  - $\geq$1 pixel hit.
  - $\geq$ 5 SCT hits.
  - Fewer than 3 Pixel or SCT holes (A hole is defined as an active sensor traversed by the track but containing no hits)
  - at least 10% of the TRT hits originally assigned to the track are included in the final fit (in the $0.1 < |\eta| < 1.9$.)
- The requirements to satisfy $p_T > 10$ GeV and $|\eta|$ <2.5 are applied as additional track selection criteria in order to be confined to the region with ID coverage.

- The significance of the transverse impact parameter and longitudinal impact parameter are imposed to have the absolute value of $|d_0/\sigma_{d_0}|$ <3 and $|z_0 sin\theta|$ <0.5 mm.

Four muon identification selections (*Medium*, *Loose*, *Tight*, and *High-p_T*) have been produced with increasing background rejection. In this dissertation muons require to satisfy the following medium quality which minimizes the systematic uncertainties associated with muon reconstruction and calibration [72] :

- hits in at least two MDT layers except for tracks in the $|\eta| < 0.1$ region, where tracks with at least one MDT layer but no more than one MDT hole layer are allowed.
- q/p significance < 7.

Similarly to electrons, the prompt muons have to be isolated from the heavy-falvour hadron semi-leptonic decays. Consequently, the same isolation variables as for electrons are available. The track isolation variable $p_T^{varcone30}$ is the sum of the transverse momenta of the tracks with $p_T > 1$ GeV with a cone of size $\Delta R = \min\Big(10 \text{ GeV}/p_T^{\mu}, 0.3\Big)$. However, if there is a jet within ΔR <0.4 of the muon which has fewer than three associated tracks, the jet is removed and the muon is kept to decrease the removal of high energy muons which lose significant energy in the calorimeter. The calorimeter-based isolation variable, $E_T^{topcone20}$ represents the transverse energy of topological clusters within a cone of size ΔR=0.2 around the muon. The Medium combined muons also require to satisfy the "Gradient" isolation working point as per electrons. The highest p_T muon for both 2015 and 2016 sets require to be $p_T >$ 27 GeV.

5.4 Jets

Quarks and gluons produced in the hard interactions can not be isolated which is a consequences of color confinement. Furthermore, they spread into a spray of particles in a process called *hadronization*. To identify a possible seed quark of a specific particle, a jet is defined as a grouping of the particles produced in the hadronization.

Different categories of jets can be defined and differ in the algorithm used to combine and build a jet (for example, both algorithm type and jet size parameter), the

type of input objects to the jet algorithm and the energy scale calibration. The standard jet identification algorithm used by ATLAS collaboration is the anti-k_t algorithm [73] with a R-parameter of 0.4. A variety of objects can be used as inputs to this algorithm: inner detector tracks [74], calorimeter energy deposits [75] or a combination of both [76]. The majority of ATLAS analyses use jet reconstruction based on calorimeter deposits, which is called calorimeter jets [77]. The analysis of this dissertation follows the ICHEP procedure [8] which uses the calorimeter jets. Therefore, the candidate jets are reconstructed from three-dimensional topological energy clusters so-called topo-clusters [78] in the calorimeter. The clustering algorithm starts by selecting the cells with energy deposits above 4σ, where σ is the noise standard deviation in the calorimeter cells, and iteratively added all the neighbor cells with energy above 2σ. Then, the energy of each topological cluster is measured at electromagnetic (EM) scale which correctly measures the energy deposited by electromagnetic showers in the calorimeter.

The reconstructed jets are then calibrated to correct the energy of the reconstructed jets in the detector to correspond to one of the truth particle jets. this is recover by the application of a jet energy scale (JES) derived from MC simulation. It corrects for detector effects due to the mismeasurement of the deposited energy, the energy lost in active regions of the detector and the difference in calorimeter granularity. The jet energy response is the ratio between the energy measured in the reconstructed jets, E^j_{EM}, and the particle-jet energy, E^j_{truth}. After the jet calibration procedure, the data-to-MC differences are assessed using *in-situ* techniques based on 13 TeV data [79]. This correction is only applied for jets in data, to restore the energy of the jets reconstructed in data to that from the MC simulation.

After energy calibration, jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Then, a jet cleaning procedure is imposed to remove events with "bad jets", whereby not all the jets which are reconstructed in detector have their origin in the pp collisions. For example, showers induced by cosmic rays or beam-gas interactions. The selection to identify the fake jet in the analysis is called the *BadLoose* selection [80–82]. It is the largest possible proportion of events with fake jets. A jet is identified as a BadLoose jet if it satisfies at least one of the following criteria:

1. f_{HEC} >0.5 and $|f_Q^{HEC}|$ >0.5 and $\langle Q \rangle$ >0.8:

   $\langle Q \rangle$ is the jet quality factor which is defined as the energy-squared weighted average of the shape of the electrical signals (pulse shape) collected in every calorimeter cell. The normalized quality of this variable is $0 < Q < 1$. If the jet has a significant deviation from the quality factor, it is rejected. The

$f_{HEC} >$ is the fraction of the energy in the hadronic end-cap calorimeter HEC. If there is large noise in the HEC and the fraction of the jets energy in the HEC exceed 50%, the jet could be bad. The $|f_Q^{HEC}|$ is the fraction of the energy in the HEC calorimeter cells of a jet with poor signal shape quality defined as Q_{cell}^{LAr} >4000.

2. $|E_{neg}|$ >60 GeV:

   $|E_{neg}| >$ is the negative energy in the cells arising from electronic noise or out-of-time pile-up.

3. f_{EM} >0.95, f_Q^{LAr} >0.8, $\langle Q \rangle$ >0.8 and $|\eta|$ <2.8:

   f_{EM} is the fraction of the jet energy deposited in the electromagnetic calorimeter which must be between 5% and 95% to reduce the noise effects from the non-collision backgrounds and EM calorimeter it self. The f_Q^{LAr} is the fraction of energy corresponding to LAr cells of a jet with poor signal shape quality defined as Q_{cell}^{LAr} >4000.

4. f_{max} >0.99 and $|\eta|$ <2:

   f_{max} maximum energy fraction in one calorimeter layer.

5. f_{max} >0.99, f_{ch} <0.05 and $|\eta| < 2$:

   f_{ch} is the jet charge fraction ratio of $\frac{\sum p_T^{\text{tracks}}}{\text{jet}p_T(\text{calib})}$

6. f_{EM} <0.05 and $|\eta| \geq 2$.

To reduce contamination due to the hard QCD jets originating from a pile-up vertex which are a real jets, they must be tagged and rejected using the *jet vertex fraction* (JVF) variable [83]. The JVF is a variable used in ATLAS to identify the primary vertex from which the jet originated. To remove jets which are not associated with hard-scatter *primary vertex* (PV), the cut on JVF variable can help. By using tracks reconstructed from the ID information and assigning tracks to jets adding infinitesimal p_T's (p_T >500 MeV) to the jets which are considered in the JVF calculation. Then, the JVF is defined as the ratio of the scalar sum of the p_T of matched tracks that originate from a given PV to the scalar sum of p_T of all matched tracks in the jet, independently of their origin that the JVF is defined for each jet i with respect to primary vertex j, PV_j.

$$\text{JVF}(\text{jet}_i, \text{PV}_j) = \frac{\sum_m p_T(\text{track}_m^{\text{jet}_i}, \text{PV}_j)}{\sum_n \sum_l p_T(\text{track}_l^{\text{jet}_i}, \text{PV}_n)} \tag{5.1}$$

where m runs over all tracks originating from PV_j matched to the selected jet_i, n over all primary vertices in the event and l over all tracks originating from PV_n matched to jet_i. The JVF is chosen to be JVF>0.59 for both data and simulation and jets with p_T <60 GeV and $|\eta|$ <2.4 are rejected if they have a low value of the JVF output. Figure 5.2 shows an example of a jet with high JVF originating from the real primary vertex PV_1 $\left(\frac{\sum_{m=1}^{4} p_T(\mathrm{track}_m^{jet_a}, \mathrm{PV}_1)}{\sum_{n=1,2}\sum_{\substack{l=1,4 \\ l=5}} p_T(\mathrm{track}_l^{\mathrm{jeta}}, \mathrm{PV}_{n=1,2})}\right)$.

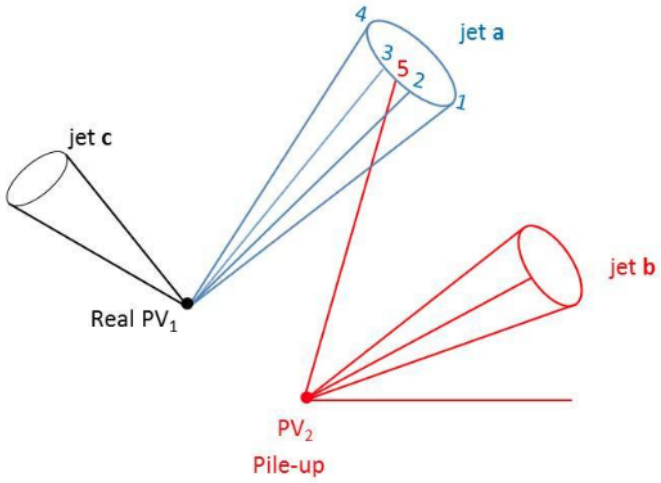


Figure 5.2: Schematic representation of the jet vertex fraction JVF principle.

During jet reconstruction, no distinction is made between the identification of electrons and jets. Therefore, to avoid double-counting of electrons as jets, the jet lying within ΔR=0.2 of selected electrons are rejected. After this, electrons which are within ΔR=0.4 of a remaining jet are discarded [8].

5.5 b-tagging

One of the most important ingredients of the ATLAS physics program and are considered a crucial task in the analysis presented in this dissertation, is the ability to identify experimentally the flavour of a jets separating b-jets from c- and light-flavour parton jets. The b-tagging procedure relies upon results from the fragmentation of b-quarks characterized by its long life-time $\tau \sim 1.5$ ps, b-hadrons can travel about 3 mm in the transverse direction and then typically decaying at a sufficiently large distance from the *primary vertex* (PV). Therefore, displaced secondary decay vertices of this b-hadron can be reconstructed, see Figure 5.1. The transverse impact parameter, which is the distance of the closest approach of the background extrapolation of the track to the PV in the $r - \phi$ plane, tends to have larger value

for the tracks from b-hadron decay products. This can be exploited to distinguish from tracks originating from the PV. The longitudinal impact parameter, z_0, is the difference between the z coordinates of the position and closest track way to that vertex.

Several algorithms have been developed in ATLAS to perform the b-tagging of jets which utilize the previous properties. Two distinct basic b-tagging algorithm categories are used in ATLAS [84]:

- Impact parameter based algorithm:

  IP3D; This is a likelihood-based combination of the transverse and longitudinal impact parameter significance. Input variables are the probability density functions (PDFs) for the signs of the transverse and longitudinal impact parameter tracks to define ratios of the b- and light-flavour jet hypotheses.

- Displaced vertices based algorithms:

  SV algorithm: This attempts to reconstruct an inclusive SV based on likelihood combination technique for various variables such as the vertex mass of all the charged particle tracks used to reconstruct the vertex, the ratio of the sum of energies of the previous tracks to the sum of the energies of all tracks in the jet and the number of two-track vertices.

  JetFitter; relies on the long life-time of b and c hadrons and on the full decay topology of the b-hadron inside the jet by finding the common line of the primary vertex, and the secondary and tertiary vertices of charm decay vertices (called cascade, $b \to c(\bar{c}) \to \mu^+(\mu^-)$).

The algorithm used in the present analysis, being one of the most common and powerful ones in ATLAS, is named *MultiVariate Algorithm 2* (MV2c20). It uses the outputs of the IP3D , SV and JetFitter as inputs to a *boosed decision tree* (BDT) approach (explained in 4.2.1) to discriminate b-jets from light (u, d, s-quark or gluon jet) and c-jets. The algorithm is trained on b-jets as signal and a mixture of 80% light-jets and 20% c-jets as background. The training is performed using simulated dileptonic $t\bar{t}$ events. The selected $t\bar{t}$ sample has a very high branching ratio of b-jets, BR($t \to Wb$) is about 99.8%, and both W bosons are forced by selection criteria to decay into leptons. There are several cut values applied to the output distribution of the MV2c20 variable to provide a specific b-jet efficiency on a $t\bar{t}$ sample. The 77% b-jet efficiency operating point is chosen in this analysis which requires a jets to have an MV2c20 output value greater than -0.4434. The MV2c20 output distribution is shown in Figure 5.3.

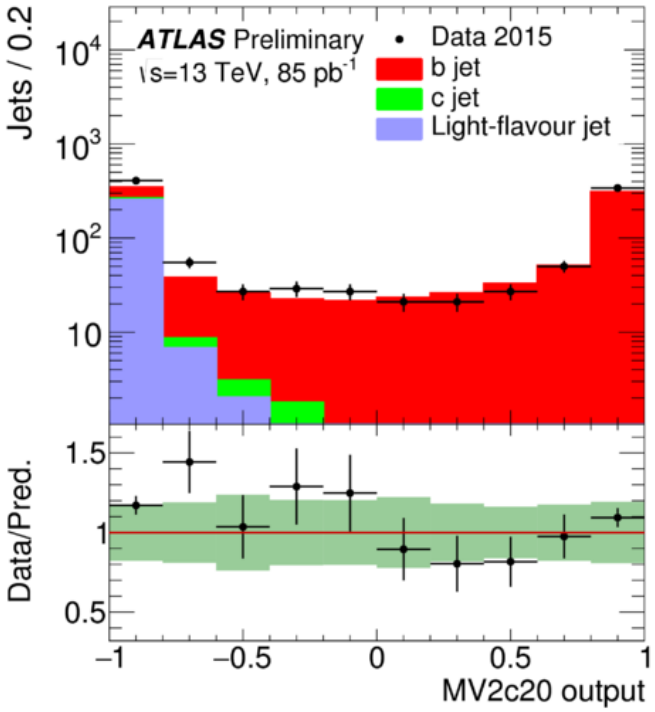


Figure 5.3: Illustration of the output distribution of the MV2c20 algorithm applied to jets from the $t\bar{t}$ dominated $e\mu$ sample.

5.6 Missing transverse energy

Particles like neutrinos, new weakly-interacting particles and the contributions of additional interactions of partons in the same bunch crossing or out-of-time collisions (for example, collision from a different bunch crossing which leaves tracks or deposites energy which are included in the events) called pile-up can lead to the momentum imbalance in the plane transverse to the beam axis. The missing transverse momentum in the event is defined as the magnitude of the negative vector sum p_T of all selected and calibrated physics objects in the event denoted as missing transverse energy, E_T^{miss} [85, 86].

The calorimeter energy deposits are associated and identified with high-p_T parent objects in a specific order to avoid the double counting of the same object. It is calculated starting from electrons, photons, hadronically decaying $\tau-$leptons, jets and finally muons. Therefore, the E_T^{miss} is the deposits energy not associated with any of the mentioned objects:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,soft} + E_{x(y)}^{miss,\mu} \tag{5.2}$$

The term $E_{x(y)}^{miss,soft}$ includes the remaining energy deposits not associated with the

high-p_T physics objects which is most affected by pile-up. Those signals associated with the high-p_T physics objects in E_T^{miss} are referred to as "hard terms". There are several algorithms which can be used to reconstruct the $E_{x(y)}^{miss,soft}$ term using either calorimeter energy depoits, tracks or both [86]. The main algorithm for the soft term reconstruction used by ICHEP [8] is also used by this analysis. It fully relies on tracks, the so-called *Track Soft Term* (TST). This algorithm uses only the ID tracks not associated with high-p_T physics objects. This algorithm is characterized by a better response to the amount of the pile-up. It removes almost all of the in-time pile-up dependence by allowing excellent vertex matching for the soft term. The missing transverse momentum is not used for event selection but it is used in the event reconstruction.

Chapter 6

Data and Monte Carlo samples

6.1 ATLAS Data Sample

The analysis described in this thesis is based on pp collisions data at $\sqrt{s}$=13 TeV collected by the ATLAS detector using the full 25 ns 2016 (150 runs) data corresponding to an integrated luminosity of 32.9 fb^{-1} $\pm$ 2.2%, see [61, 87]. The final integrated luminosity is sub-divided into several periods which is also divided into runs corresponding to a shorter interval of data-taking. In general, a total integrated luminosity was recorded when the ATLAS sub-detector high-voltage (HV) is ramped up, and the stable beam flag is set. When ATLAS achieves stable conditions, it begins recording luminosity, see Figure 6.1 (yellow). Luminosity delivered by LHC is also shown (green), the difference between recorded and delivered luminosity is largely due to Trigger and Data Acquisition system (TDAQ) inefficiency and the previously mentioned ramp-up time of the HV systems.

The analysis in this thesis, which is the $t\bar{t}H$ channel with a single lepton originating from a leptonic decay of a W, requires events to be accepted by at least one electron or one muon trigger [88, 89]. Events are selected using different single lepton triggers with different p_T thresholds, which are then combined in a logical "OR" in order to increase the overall efficiency. The triggers listed in Table 6.1 were unprescaled up to data-taking period D4 in 2016. The three electron triggers require lepton to have p_T >24, 60 and 140 GeV respectively, and in addition, there is an isolation requirement for the low-p_T trigger. Similarly, for the two muon triggers, the low-p_T trigger also has an isolation requirement, and the high-p_T trigger has a threshold of 50 GeV. An offline analysis cut of 25 GeV is applied to both electron and muon for this running period. The additional unprescaled triggers listed in Table 6.2 were

applied starting from period D4, where the lower p_T threshold was increased to p_T >26 GeV, and the other triggers were kept. The offline analysis cut applied was p_T >27 GeV.

e+jets	HLT_e24_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0
μ+jets	HLT_mu24_ivarmedium OR HLT_mu50

Table 6.1: The unprescalled triggers up to period D4 of data taking in 2016 with the lowest lepton p_T threshold of 24 GeV.

e+jets	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e140_lhloose_nod0
μ+jets	HLT_mu26_ivarmedium OR HLT_mu50

Table 6.2: The unprescalled triggers starting from period D4 of data taking in 2016 with the lowest lepton p_T threshold of 26 GeV.

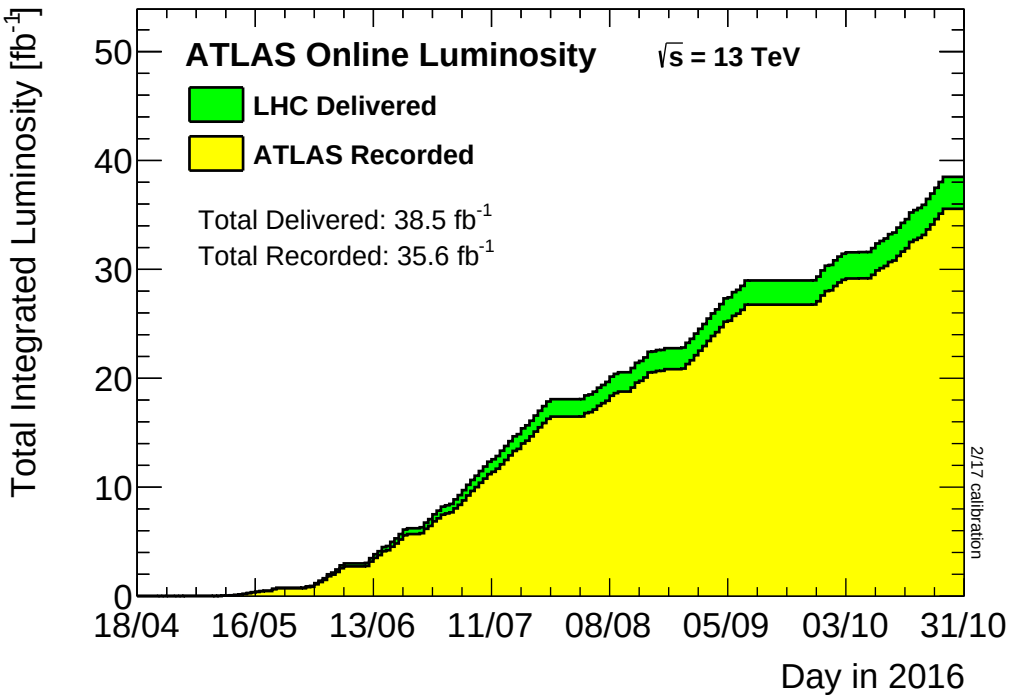


Figure 6.1: Cumulative luminosity versus time delivered to (green) and recorded by ATLAS (yellow) as seen by the ATLAS detector during the 2016 stable beam runs in $\sqrt{s}$ =13 TeV pp collisions.

6.2 Signal and Background Modeling

After preselection, the $t\bar{t}$+jets production is the main background to this analysis (semi-leptonic process) even though some other processes have larger cross-section because a large fraction of the other backgrounds can be removed depending on the criteria mentioned in 7.3. The $t\bar{t}$+jets is divided into an irreducible and a reducible fractions. The irreducible $t\bar{t} + b\bar{b}$ background originate from QCD [90] or EW [91] processes resulting in the same final state signature as signal. The reducible fraction corresponds to background $t\bar{t}$ associated with c-quark jet referred to as $t\bar{t} + c\bar{c}$ and can mimic signal when the c-quark jets are misidentify as b-quark jets. The other reducible background arises from $t\bar{t}$+light jets when the $t\bar{t}$ is produced in association with light quark (u, d and s) or gluon jets. The other background contributions originate from single top quark production, the production of W or Z bosons in association with jets (V+jets), diboson (WW, WZ, ZZ) production, as well as from the associated production of a vector boson and $t\bar{t}$ pair ($t\bar{t} + V$). After preselection, the multijet background contributes only a very small fraction. It falls under the scope of this analysis because the light jets or photons are misidentified as electrons, or the events contain a non-prompt lepton. Therefore, almost all of the multijet background is removed by the lepton cut, and after requiring two or more b-tagged jets reduces this even further. In this chapter, a description of the simulation of each background and of the signal is detailed. The Higgs boson mass and the top-quark mass are set to 125 GeV and 172.5 GeV, respectively, in all the simulated samples. Some of the relevant parameters of the MC samples used in the analyis are summarized in Table 6.3.

Sample	Generator	PDF	Shower	Normalisation
$t\bar{t}H$	MG5_aMC [92]	NNPDF3.0NLO	Pythia 8.2 [93]	(N)NLO
$t\bar{t} + jets$	PowHeg [94–96]	NNPDF3.0	Pythia 8.2 [97]	NNLO+NNLL
W+jets	Sherpa [98]	CT10	Sherpa 2.2.1	NNLO
Z+jets	Sherpa	CT10	Sherpa 2.2.1	NNLO
Single top (s-channel, Wt)	PowHeg	CT10	Pythia 8.2	aNNLO
Single top (t-channel)	PowHeg	CT10f4	Pythia 8.2	aNNLO
$t\bar{t}V$	MG5_aMC	NNPDF3.0NLO	Pythia 8.2	NLO
Diboson	Sherpa	CT10	Sherpa 2.1.1	NLO

Table 6.3: List of generators and parameters used to simulate various processes.

6.2.1 $t\bar{t}$+jets production

Because $t\bar{t}$ is the dominant background in the $t\bar{t}H$ semi-leptonic analysis, there is a requirement for a $t\bar{t}$ simulation which describes correctly the final state for emission of additional jets and the heavy flavour fraction. In the analysis, the simulation should model the full final state correctly in parallel with the normalization.

The nominal $t\bar{t}$ background is generated using the Powheg-Box x v2 NLO generator with the NNPDF3.0 parton distribution function (PDF) set [99]. It is interfaced to Pythia 8.2 with appropriate A14 tune to model the parton shower and hadronization. The *hdamp* parameter in Powheg, controls the p_T threshold of the first additional emission beyond the Born approximation. Therefore, hdamp regulates the p_T of the recoil emission against the $t\bar{t}$ system. It is set to 1.5 times the top quark mass m_t=172.5 GeV. The previous setting of Pythia and hdamp was different from the one that is used in thsi analysis and ICHEP [61]), where the hdamp was set to top quark mass and Pythia 6.428 was used.

The bottom and charm hadron decays are simulated by the EvtGen v1.2.0 [100] program. The sample is normalized to the top++2.0 [101] theoretical cross-section of 832^{+46}_{-51} pb, calculated at next-to-next-to-leading order (NNLO) in QCD and including resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms [102–106].

The $t\bar{t}$+jets sample is generated inclusively but categorized according to the flavour of partons that match the additional jets those not originating from the decay of the $t\bar{t}$ system. Therefore, the events are classified into three orthogonal samples; $t\bar{t}$+light jets, $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$. Particle jets are reconstructed and classification based on an anti-k_t algorithm with a radius of R=0.4. This algorithm does the matching of hadrons to particle jets which are built from stable truth particles with p_T >15 GeV and $|\eta|$ <2.5 excluding muons and neutrinos. Firstly, the particle jet is labeled b- or c-jet if it achieves the truth matching which considers the set of b- and c-hadrons with p_T >5 GeV not originating from $t\bar{t}$ decay products. This excludes hadrons from b-quarks from top decays and the hadrons produced by c-quarks from hadronic W boson decays. If the particle jet is matched within ΔR <0.3 to a b- or c-hadron, then it is labeled as a b- or c-jet. Events with at least one b-jet not originated from top decay products are labeled as $t\bar{t}+ \geq 1b$ events. Events which do not fulfill this criterion, but include at least one c-jet not from a W decay are labeled as $t\bar{t}+ \geq 1c$ events. The set of $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ events are generically reffered to as $t\bar{t}$+HF events, HF denoting "heavy flavour". The remaining events

with no heavy flavour jets are labeled as $t\bar{t}$+light jets events. For the purpose of using these samples samples to estimate the systematic uncertainties, there is a finer categorization. Events with exactly two b- or c-jets are labeled $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$, respectively. Events that have single b- or c-jet matched a single b- or c-hadron are referred to as $t\bar{t}+b$ or $t\bar{t}+c$ respectively. Those events with a single particle jet matched to a b-hadron or c-hadron pair are labeled $t\bar{t}+B$ and $t\bar{t}+C$, respectively.

$t\bar{t}+\geq 1b$ modeling

The main irreducible background in signal regions is the production of a pair of top quarks in association with a bottom quark-antiquark pair ($t\bar{t}+b\bar{b}$) which plays a key role in the regions of high sensitivity to the $t\bar{t}H$ signal (the signal that have high jet and in particular b-jet multiplicities). Therefore, it is important to have precise modeling of this background. There is a further improvement whch is made to the modeling of $t\bar{t}+\geq 1b$ by reweighting the Powheg+Pythia 8.2 prediction to the NLO production from the fully matched NLO predictions (generated, showering and hadronization) with massive b-quarks within the Sherpa 2.1.1 interfaced with OpenLoops (referred to as SherpaOL in the following) framework [95, 107–110] using the CT10 PDF set. This reweighting can reduce the perturbative uncertainties on the cross-section from 70-80% of the LO calculation, down to 15-20% [111, 112]. Moreover, the reweighting is performed for different topologies of $t\bar{t}+b\bar{b}$ ($t\bar{t}+b\bar{b}$, $t\bar{t}+b$, $t\bar{t}+B$, etc). The inclusive $t\bar{t}+b\bar{b}$ cross-section is kept constant throughout all the reweightings while the relative cross-section in each category is adjusted to the NLO prediction. There is an exception for two of the $t\bar{t}+b\bar{b}$ contributions which are not included in the NLO prediction by SherpaOL. Contributions from events where $b\bar{b}$ pairs are produced via *multiple parton interaction* (MPI) and overlaying a $t\bar{t}+jets$ events, which is labeled as "$t\bar{t}$+MPI", or via gluon radiation from the top decay products labeled as *final state radiation* "$t\bar{t}$+FSR". These two contributions, MPI and FSR, see Figure (6.2), have to be identified and excluded from the comparison to the SherpaOL sample. The ICHEP results [61] show the relative contribution of the different $t\bar{t}+b\bar{b}$ event categories to the total $t\bar{t}+b\bar{b}$ cross-section at generator level. It demonstrates that Powheg+Pythia 8.2 is able to reproduce reasonably well the $t\bar{t}+\geq 1b$ content of the NLO SherpaOL prediction. Also, it shows that SherpaOL predicts a higher contribution in the $t\bar{t}+B$ category, as well as in $t\bar{t}+\geq 3b$. The two contributions, MPI and FSR, have to be identified and excluded from the comparison between the SherpaOL sample, which does not include MPI and FSR contributions, to the Powheg+Pythia sample that includes these two contributions. To improve the

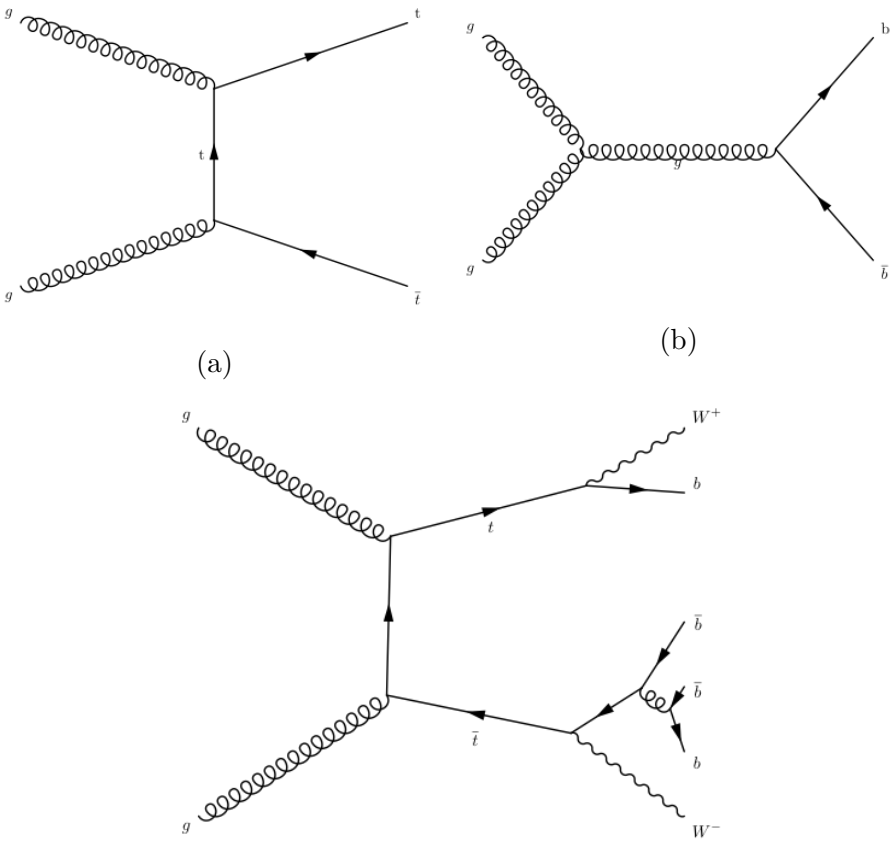


Figure 6.2: Example Feynman diagrams for the $t\bar{t}$+PMI and $t\bar{t}$+FSR and contributions.

agreement of the different variables in each category, the reweighting is done at the generator level using several kinematic variables [8]. The first reweighting is based on the p_T of the top and $t\bar{t}$ system. In the event topology with only one additional heavy flavour jet, p_T and η of that jet are used as a reweighting variables. A final reweighting is performed using ΔR and p_T of the dijet system of the topologies with two or more heavy flavour jets.

$t\bar{t}+ \geq 1c$ jets modeling

The $t\bar{t} + c\bar{c}$ background is considered the second largest background in the signal regions. In the case of $t\bar{t}+ \geq 1c$ production, it is generated using the inclusive Powheg+Pythia $t\bar{t}$ sample where the $t\bar{t} + c\bar{c}$ background events is produced in parton shower. To achieve more confidence in the prediction, a novel simulation using MadGraph5_aMC@NLO (MG5_aMC) and Herwig++ to generate $t\bar{t} + c\bar{c}$ in the matrix element at NLO in three flavour scheme (3FS) using the CT103f PDF set. The

systematic uncertainty is assessed by calculating the difference between the sample produced with $t\bar{t} + c\bar{c}$ in the matrix and with the inclusive $t\bar{t}$. The reweighting is applied to the Powheg+Pythia sample in a similar way as for $t\bar{t} + b\bar{b}$ to preserve the overall $t\bar{t} + b\bar{b}$ normalization. This study is detailed in ref [113].

6.2.2 W/Z+jets production

The background contribution from the production of W/Z with additional jets, and diboson production in association with jets are simulated with Sherpa 2.2.1 generator. The matrix element calculation is performed using up to four partons at LO and up to two partons at NLO using Comix [114] and OpenLoops matrix element generators which are then merged with Sherpa parton shower [108] using the ME+PS@NLO prescription [115]. The PDF set used in the calculation is CT10 with dedicated parton shower tuning developed by the Sherpa authors. The samples W/Z+jets are generated inclusively the Z+jets was split into three components according to the number of heavy flavour (HF) jets produced with the Z boson. These separate samples are; Z+light jets, $W/Z+ \geq 1b$ and $W/Z+ \geq 1c$. Both W+jets and Z+jets processes are normalized to the NNLO cross-section [116]. By following the same approach, the diboson+jets samples are generated but with up to one (ZZ) or zero (WW, WZ) additional partons at LO.

6.2.3 Multijet Background

The selection of events of semi-lepotonic decay mode required one charged lepton in the final state from the W boson decay, referred to as "prompt" "real" leptons. Appropriate cut criteria are applied to select these leptons, see Sections 5.2 and 5.3. However, there are some signals which fulfill these selection criteria, giving rise to so called "non-prompt" or "fake" leptons which are a particular problem in multijet background. In the case of the electron channel, the multijet background consists of "non-prompt" electrons arising from photon conversions and heavy-hadron decays, or mis-identified jets with a high fraction of their energy deposited in the EM calorimeter called "fake" electrons. Muons can be labeled as "non-prompt or "fake" leptons when they originate from charged hadron decays (b- and/or c-quark) in hadronic showers, from semileptonic decays of b- and c-quarks, or from "punch-through" particles emerging from high-energy hadronic showers.

While the probability for multijet events to pass the event selection is very low, the

production cross-section of multijet events is orders of magnitude above $t\bar{t}$ production. This background is very difficult to model accurately with MC because there are *many* ways to produce it. Hence, the better way is to estimate the corresponding yield for this background via data-driven method employing the so-called *Matrix Method* (MM) [117]. The inputs of the MM are flat tuples from the top software package AnalysisTop-2.3.41 and TopFakesUtils-00-00-08 run over TOPQ1 derivations. Depending only on the different lepton identification criteria, two samples are defined. A "tight" sample where the lepton satisfies the identification criteria described previously in sections 5.2 and 5.3. As for the "loose" sample, some of the lepton identification or isolation requirements are excluded. In this method the tight selection is a subset of the loose selection.

In a data sample containing events with a single lepton, the number of events with one tight lepton (N^t) and the number of events with one loose lepton (N^l) can be expressed as a linear combination of the number of events with real and non-prompt and fake leptons

$$N^t = N^t_{\text{real}} + N^t_{\text{fake}}, \tag{6.1}$$

$$N^l = N^l_{\text{real}} + N^l_{\text{fake}} = N^t_{\text{real}}/\epsilon_{\text{real}} + N^t_{\text{fake}}/\epsilon_{\text{fake}}, \tag{6.2}$$

where $\epsilon_{\text{real}} = \frac{N^t_{\text{real}}}{N^l_{\text{real}}}\left(\epsilon_{\text{fake}} = \frac{N^t_{\text{fake}}}{N^l_{\text{fake}}}\right)$ is the probability of real (non-prompt and fake) leptons in the loose selection that pass the tight selection and both measured in data control samples in either real or non-prompt or fake lepton. Assuming that ϵ_{real} and ϵ_{fake} are known quantities, then equation 6.2 can be solved for the fake background N^t_{fake} as follows:

$$N^t_{\text{fake}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}}(\epsilon_{\text{real}} N^l - N^t) \tag{6.3}$$

The matrix method efficiencies ϵ_{real} and ϵ_{rmfake} depend on the kinematic properties of the event, such as the number of jets or b-jets. To correctly account for this, an event weight is computed from the efficiencies,

$$w_i = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}}(\epsilon_{\text{real}} - \delta_i), \tag{6.4}$$

where δ_i equals unity if the loose event i passes the tight event selection and 0 otherwise.

6.2.4 Measurement of the real efficiency

The real efficiencies ϵ_{real} are determined in data using the *tag-and-probe* method from the Z boson decay $Z \to ee$ or $Z \to \mu\mu$ in the control regions. Events with a pair of opposite-sign loose or tight leptons and at least one jet are selected. The lepton is selected as "probe" if it passes the loose criteria and the second electron is selected as a "tag" if it passes the tight criteria. Table 6.4 shows the two levels of lepton selection requirements.

$$\epsilon_{\text{real}} = \frac{N^t_{\text{real}} \text{ (tags)}}{N^l_{\text{real}} \text{ (probes)}} \tag{6.5}$$

The best method to extract ϵ_{real} in each considered bin is the *side-band* method. This method requires the invariant Z boson mass of the dilepton opposite-sign-charges system to be between 60 and 120 GeV, which is considered as a signal. The background is considered to be the invariant Z boson mass distribution of the same-sign lepton pair, which has a linear shape over the whole mass range from side-band to side-band.

The number of background events are estimated by extrapolating the side-band distributions and subtracting from the signal region.

6.2.5 Measurement of the fake efficiency

The fake efficiency ϵ_{fake} is measured in dedicated fake enriched control regions. It is determined as the difference between data and MC ($t\bar{t}$, single top, diboson, W+jets and Z+jets) in the tight selection divided by the difference in the loose selection, for each control region. Events were selected as follows [8]:

- exactly one loose or tight lepton
- at least one jet
- Control region used in the e+jets channel: E_T^{miss} <20 GeV
- Control region used in the μ+jets channel: muon d_0 >5
- Events are required to pass the single lepton pre-scaled triggers:
  HLT_e24_lhmeduim_nod0_L1EM18VH and HLT_mu24

In the multijet background, the real lepton events contribute with a proportion that can not be neglected. Consequently, this contribution has to be subtracted from the loose and tight samples in each control region

$$\epsilon_{\text{fake}} = \frac{N^t_{\text{fake}}}{N^l_{\text{fake}}} = \frac{N^t - N^t_{\text{real}}}{N^l - N^l_{\text{real}}}. \tag{6.6}$$

5

	Loose selection	Tight selection
Electron identification level	LooseAndBLayerLH	TightLH
Muon identification level	Loose	Medium
Lepton isolation requirement	Loose	Gradient

Table 6.4: Summary of the lepton selection requirements.

6.2.6 Other Simulated Backgrounds

There are other backgrounds which are contribute with a small cross-sections which becomes negligible in the signal regions. However, their absence in the control regions could affect the result of the fit which would propagate to the signal region which means there is a requirement to consider all these small backgrounds for completeness.

Single top production is considered an important contribution to the background. It considers all the three production mechanisms; s-channel, t-channel and Wt production. The events in the t-channel are generated using the Powheg-Box v1 generator that uses the four-flavour number scheme (4FNS) for the NLO matrix elements calculations and the fixed four-flavour PDF set CT10f4. For this process, the top quarks are decayed using MadSpin [118], preserving all spin correlations. Samples of Wt and s-channel are generated with Powheg-Box 2.0 at NLO using the CT10 PDF set. The parton shower, hadronisation and the underlying events are modeled using Pythia 8.2 with the Perugia 2012 [119]. The single top quark samples are normalized to the approximate NNLO theoretical cross-section [120–122]. Overlaps between the $t\bar{t}$ and Wt final state starting at NLO is handled using the diagram removal scheme [123].

The $t\bar{t} + V$ production can be mismatched for $t\bar{t}H$ processes especially when the irreducible $t\bar{t} + Z$ background, with subsequent $Z \to b\bar{b}$ (15%) is considered. However, this will not affect very much the signal sensitivity, because its contribution is reduced to less than half of the $t\bar{t}H(H \to b\bar{b})$ signal. The $t\bar{t} + V$ events have been

generated at NLO (What is the different between generated at NLO and normalized to their NLO cross-sections) using MG5_aMC interfaced to Pythia 8.2 and the A14 [93] UE tune, using the NNPDF23LO [124] PDF set to perform the samples showering.

Samples of single top quark plus a single Higs boson are negligible in the Standard Model and not included in the analysis in this thesis because of their very low cross-section and then they will not much impact on the analysis.

6.2.7 Signal Modeling

The $t\bar{t}H(H \to bb)$ signal process is modeled at NLO accuracy using matrix elements from the MG5_aMC, with Higgs boson mass set to m_H=125 GeV. In this case Pythia 8.2 with A14 NNPDF23LO UE tune serves as an interface to shower the generated events which are inclusive in Higgs boson decays. The $t\bar{t}H$ sample is produced using the NNPDF3.0NLO [99] PDF set and factorisation (μ_F) and renormalisation scales set to $\mu_F = \mu_R = H_T/2$, where H_T is defined as the scalar sum of the transverse masses $\sqrt{p_T^2 + m^2}$ of all final state particles. The sample is normalised to the NNLO theoretical cross-section and Higgs boson decay branching fractions are taken from NLO theoretical calculations where all Higgs boson decay modes are considered [125–133].

The analysis in this thesis includes three Monte Carlo simulated samples of events. These three samples are divided depending on the $t\bar{t}$ system decay: the *dilepton* events (where $t\bar{t}$ decays leptonically) the l+jets events (where $t\bar{t}$ decays *semi-leptonically*) and the *all-hadronic* (where there are no leptons in the final state).

Chapter 7

Implementation of a kinematic fit using the KLFitter package

The *Kinematic Likelihood Fitter* (KLFitter) is a package which uses a likelihood approach written in C++ [134]. The maximization of the likelihood is performed using the Bayesian Analysis Toolkit (BAT) [60]. The KLFitter package was originally developed for the case of top quark reconstruction. However, it was designed such that it can be adapted to fit other processes. In this chapter, the implementation of a kinematic fit to the $t\bar{t}H(H \to bb)$ channel topology within the framework of the KLFitter package is described, taking into account the main aims of kinematic fitting, which are:

- to provide a test for the fit hypothesis which correctly assigns the measured objects to the final-state particles comprising the hypothesis;
- to provide improved estimates for those observables which are then subject to constraints in the case that the fit hypothesis is indeed true.

7.1 Basic Principle

The likelihood function is the base of the KLFitter method. It describes how likely the measurement will fit a specific model depending on some model parameters θ_i:

$$\mathcal{L}(\text{measurement}|\text{model}(\theta_i)) \tag{7.1}$$

The likelihood can join the *probability density function* (pdf) of several variables that are used to describe the physics process. If these variables are independent,

they can be split and the likelihood function becomes the product of their pdfs:

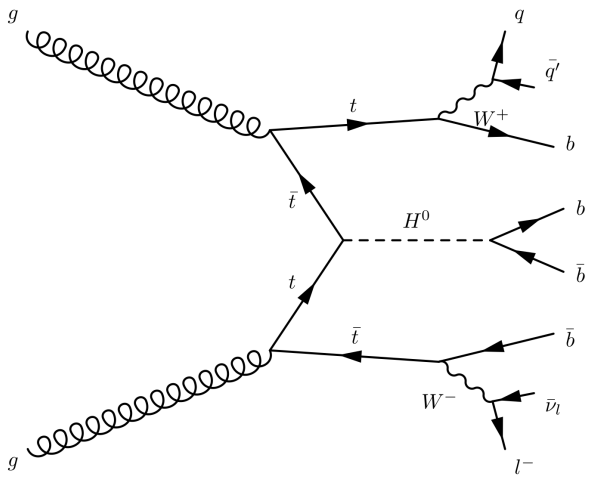
$$\mathcal{L} = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_1, ..., \theta_k) \tag{7.2}$$

One of the most important features of the likelihood usage is the *maximum likelihood estimator* (MLE) which is the value of θ that maximize the likelihood function. The MLE is usually done by taking the first derivative of the likelihood function. To simplify the MLE, the log function is used such that the derivative of the sum can be calculated rather than the derivative of the product. This concept is important to rank the permutation which is discussed in section 7.4.1.

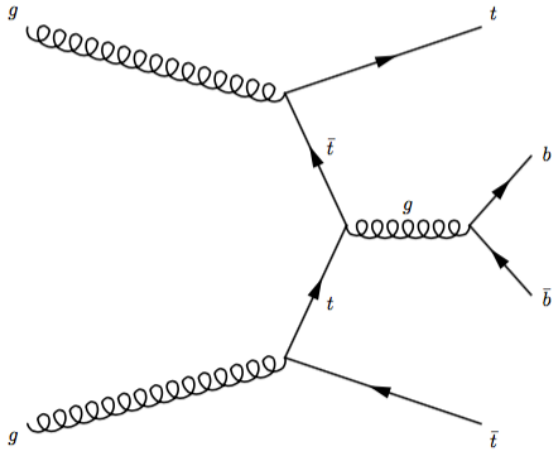
7.2 Typical application

In the present study, the full reconstruction of the top-quark four momenta is important to find the rare channel involving top quarks within $t\bar{t}H$ production. However, it is known that the top quarks has a short lifetime, hence, their properties can be studied depending on their decay mode and their corresponding signatures in the detector. In the hard-scattering process, the quarks (parton level) will hadronize to form stable particles which are referred to as jets. It is expected that a certain number of quarks will produce the same number of jets. However, it is difficult to decide which jet belongs to the initial quark, therefore, a kinematic fit is needed to find the best corresponding match between them (reconstruction level). Furthermore, the $t\bar{t}$ background and $t\bar{t}H$ signal look similar and the best way to distinguish between them is by extracting Higgs boson mass through a reconstruction algorithm, as shown in Figure 7.1.

In the current application of the KLFitter, a new likelihood function (KLFitterLH) is modified to fit the $t\bar{t}H$ mode. It is valid in the l+jets channel for production of a Higgs boson in association with a pair of top quarks. There are several KLFitter Jet selection modes associated to KLFitterLH, requiring intensive investigations to study the efficiency of each jet selection mode, are implemented and explained in Section 7.3. According to the kB6 mode which is adopted in our analysis as a result of the MVA results, see Section 9.1.1, the reconstructed event are required to have six particle jets out of which four contain b-quarks. Exactly six jets corresponds to no additional jets come from the initial state radiation **ISR** or final state radiation **FSR** and no lost jets that may not pass the preselection cut criteria, in principle, there are 720 (6!) different permutations to assign the jets to six quarks of the ttH decay system. Since both light quarks jets combinations yield the same W



(a)



(b)

Figure 7.1: $t\bar{t}H$ channel (top), shown the top quark and Higgs decay modes in the semileptonic channel and generic $t\bar{t}$ background (bottom)

candidate and the two b-quarks from the Higgs decay are indistinguishable, there are 180 (6!/4) possible permutations for the $t\bar{t}H$ decay mode. The precision of a measurement may decrease as a result of the *combinatorial background*. To improve the efficiency of the reconstructed events, the *best permutation* is chosen which is described in Section 7.4.1. The other 179 permutations are invariably wrong due to the combinatoral background.

7.3 Event Preselection

The determination of the offline event preselection for the experimental signature of a $t\bar{t}H(H \to b\bar{b})$ event in the single lepton decay topology of the $t\bar{t}$ is characterized by an isolated electron or muon with high transverse momentum, as described in chapter 5. As the reconstruction is performed using KLFitter, there are five jet selection modes available to choose from, each of which includes jets and b-tags in a specific and different order of priority, up to a limit set by the mode itself. At the start of this KLFitter study, the l+jets decay channel of top-quark pairs is used and the likelihood function is defined in the class LikelihoodTopLeptonJets, where four jet selection modes are studied (see Section 8.1). These four modes are initially named: kL4, kL5, kB4 and kB5, see Table 7.1. Subsequently, KLFitter is applied in the l+jets channel for production of a Higgs boson in association with a pair of top quarks ($t\bar{t}H$) and the likelihood function is defined in the KLFitterLH class equation 7.4. Four extra jet selection modes are tested in this KLFitter type, kL6, kB6, kB7 and kB8, see Table 7.1. Note, KLFitterLH is valid only for (jet $\geq$ 6, b $\geq$ 4) because of the use of a b-tag method called kVetoNoFitLight, see Section 7.4. These modes mean add firstly four b-tags then, adding the other light jets ordered by p_T to have 6, 7, 8 jets, respectively. The kBtagPrioritySixJets mode is adopted in this analysis as a result of the MVA studies, see Section 9.1.1. Accordingly, the single lepton $t\bar{t}H$ channel is seeded by the lowest un-prescaled single electron and single-muon triggers using the 2016 data sets described in Section 6.1. In addition, events are required to contain a certain number of jets and b-tagged jets. At least 6 jets are required satisfying the requirements of Section 5.4. For events with exactly 6 jets, at least 4 b-tagged jets are required at the 77% working point, see section 5.5.

KLFitter jet selection mode	Description	Abbreviation
kLeadingFour	Includes the highest four jets in p_T, regardless of b-tagging.	*kL4*
kLeadingFive	Same as kL4, except that the highest five jets ordered in p_T are included.	*kL5*
kLeadingSix	Same as kL4, except that the highest six jets ordered in p_T are included.	*kL6*
kBtagPriorityFourJets	Includes b-tagged jets up to the limit of four, any spare slots are filled with the highest p_T light jets.	*kB4*
kBtagPriorityFiveJets	Same as kB4, with a limit of five b-tagged jets/light jets.	*kB5*
kBtagPrioritySixJets	Same as kB4, with a limit of six b-tagged jets/light jets.	*kB6*
kBtagPrioritySevenJets	Same as kB4, with a limit of seven b-tagged jets/light jets.	*kB7*
kBtagPriorityEightJets	Same as kB4, with a limit of eight b-tagged jets/light jets.	*kB8*

Table 7.1: The several KLFitter jet selection modes that are used in the analysis.

7.4 The likelihood function for the $t\bar{t}H$ channel

Figure 2.8 shows the decay chain of the $t\bar{t}H$ channel in the lepton+jet signature, with its final state consisting of four b quarks, two light quarks, one charged lepton and one neutrino. The inputs to the KLFitter model are the reconstructed particles from track/hits/caloremeter in the detector, shown in Figure (7.2), and the parameters deriving from these objects are the following:

- reconstructed energy of six jets and one electron $\tilde{E}$.
- transverse momentum of the selected muon, $\tilde{p_T}$, since the muon leaves a negligible calorimeter deposit.
- missing transverse energy components (corresponds to the transverse momentum of the neutrino), $\tilde{E}^{miss}_{x,y}$.

- angular variables, η and ϕ. These are assumed to be measured with negligible uncertainties.

The other inputs parameters are the truth particles from hard scattering of partons in the generators which are the following:

- top pole mass and width fixed to 172.5 and 1.5 GeV, respectively. It is found that the efficiency increases when the top quark masses are fixed [134].
- W-boson pole mass and width fixed to 80.4 and 2.1 GeV, respectively.
- Higgs boson mass treated as a free parameter around its pole mass of 125.0 GeV with width fixed to 0.00351 GeV [1].
- The z component of the neutrino momentum does not have a transfer function and the neutrino does not have a measured energy. Therefore, it enters the fit via the W mass constraint as is described in Appendix A.
- The two light-quark energies, the four b-quark energies (E_{blep}, E_{bhad}, $E_{b1Higgs}$, $E_{b2Higgs}$) and the energy of the charged lepton E_l are treated as free parameters. KLFitter allows small variations, roughly the size of the estimated errors on the objects themselves which give better precision. The variation are constrained by the errors of the masses of the Higgs, W_{had}, W_{lep}, top_{had} and top_{lep}.

[1] These pole masses impose constraints on the reconstructed particles in the likelihood function. For a reconstructed mass $\tilde{m}$, the likelihood function quantifies the probability of the correspondence to a truth particles with mass m and width Γ.

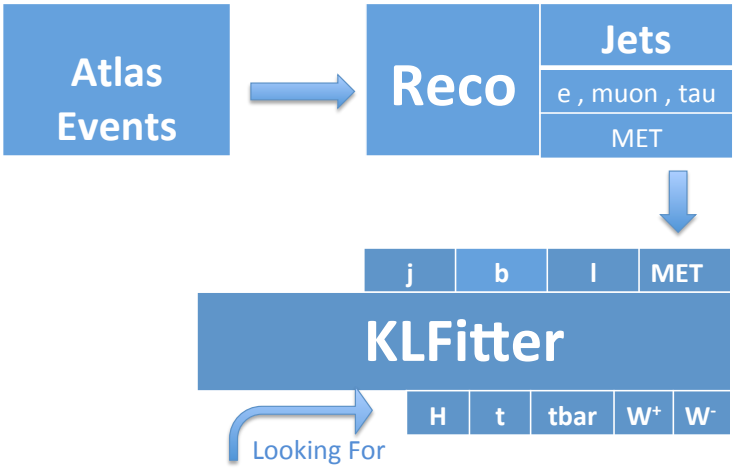


Figure 7.2: Schematic representation of the input and output parameters of KLFitter.

The constraints are applied due to the masses parametrized by the relativistic Breit-Wigner functions (BW). Further information regarding the Breit-Wigner and transfer functions is given in Sections 7.4.2 and 7.4.3. KLFitter allows small variations, roughly the size of the estimated errors on the objects themselves which give better precision e.g. m_W.

As mentioned in Section 7.1, the likelihood can be expressed as a product (convolution) of the PDFs of several input parameters. The likelihood function is composed of two parts: the transfer functions $\mathcal{L}_W$ (energy resolution of the measured jets and leptons) and the Breit-Wigner functions $\mathcal{L}_{BW}$ (functions describing the truth particle properties, masses and widths)

$$\mathcal{L} = \mathcal{L}_W * \mathcal{L}_{BW}. \tag{7.3}$$

Therefore, the overall likelihood function for the l+jets channel for production of a Higgs boson in association with a pair of top quarks ($t\bar{t}H$) is

$$\mathcal{L} = W(\tilde{E}_{jet1}|E_{bhad}) \cdot W(\tilde{E}_{jet2}|E_{blep}) \cdot W(\tilde{E}_{jet3}|E_{q1}) \cdot W(\tilde{E}_{jet4}|E_{q2}) \cdot W(\tilde{E}_{jet5}|E_{b1Higgs})$$
$$\cdot W(\tilde{E}_{jet6}|E_{b2Higgs}) \cdot W(\tilde{E}_{x}^{\mathrm{miss}}|p_{x,\nu}) \cdot W(\tilde{E}_{y}^{\mathrm{miss}}|p_{y,\nu}) \cdot \begin{cases} W(\tilde{E}_l|E_l), & \text{single electron channel} \\ W(\tilde{p}_{T,l}|p_{T,l}), & \text{single muon channel} \end{cases}$$
$$* BW\{m(q1q2)|m_W,\Gamma_W\} \cdot BW\{m(l\nu)|m_W,\Gamma_W\} \cdot BW\{m(q1q2b_{had})|m_{top},\Gamma_{top}\}$$
$$\cdot BW\{m(l\nu b_{lep})|m_{top},\Gamma_{top}\} \cdot BW\{m(b1Higgsb2Higgs)|m_{Higgs},\Gamma_{Higgs})\}. \tag{7.4}$$

The b-tagging information is including in the likelihood as a weight $\mathcal{W}_{btag}$ derived from the b-tagging efficiency ϵ and the light jet rejection $\mathcal{R}$. This weight is then multiplied with the likelihood from the fit, resulting in a quantity called the *event probability*.

$$\text{KLFitter.eventprobability} = \mathcal{L} \cdot \mathcal{W}_{btag} \tag{7.5}$$

The chosen b-tagging working point has a high efficiency ($\epsilon = 0.77$), and a high light-jet rejection rate $\mathcal{R}$=113 (1-$\frac{1}{\mathcal{R}} \approx 1$). Therefore, for an efficiency ϵ and and a rejection $\mathcal{R}$, the tagging probability of each jet in an event is

$$p_{\mathrm{btag}}(\text{tagged } |b) = \epsilon, \tag{7.6}$$
$$p_{\mathrm{btag}}(\text{not tagged } |b) = 1-\epsilon, \tag{7.7}$$
$$p_{\mathrm{btag}}(\text{tagged } |\text{light}) = \frac{1}{\mathcal{R}}, \tag{7.8}$$
$$p_{\mathrm{btag}}(\text{not tagged } |\text{light}) = 1-\frac{1}{\mathcal{R}}. \tag{7.9}$$

In the $t\bar{t}H$ signal where there are four b-jets and two light-jets, the highest weight is formed if all the jets are picked up correctly and then the weight of the event becomes around ϵ^4. In addition to the weighting method which uses the b-tag working point, there is another method available to improve the reconstruction efficiency in the KLFitter called kVetoNoFitLight. This b-tag veto mode removes the permutations before the fitting procedure if the light-jet happens to be used as a b-jet by the fit, for example if the light jet is included in the Higgs or top quarks. This specifically helps to reduce the number of permutations from 180 for six jets, to 12 permutations for four b-jets (4!/2, for two indistinguishable combinations of b-tags from Higgs plus two indistinguishable light jet combinations from W boson). This work significantly reduces the combinatoral background by removing most of the wrong permutations and also reduces the CPU/memory consumption which was a large problem for the higher KLFitter selection modes (corresponding to kB6, kB7 or kB8).

7.4.1 Ranking of the permutations

- KLFitter combines all possible meaningful associations of the measured particles to the nominal particles (truth particles) resulting in 12 permutations for 4 b-tags when the veto mode is applied in KLFitterLH.
- All permutations are fitted individually and then ranked.
- Most of these permutations (combinations) are wrong (11 permutations) and make very poor reconstructed Higgs, top quarks and W's ($H, t, \bar{t}, W^+, W^-$).
- The best ranked permutation is considered as the highest log likelihood value which represents the best assignment of jets to the heavy objects $H, t, \bar{t}, W^+, W^-$. Figure 7.3 and 7.4 show the improvement in the performance when the best permutation is picked in the $t\bar{t}$ system. The log likelihood distribution becomes narrower around the highest value of -40 and a reasonable proportion of the best permutations have an event probability between 0.8-1.0.
- There is a small but noticeable effect on the ranking of permutations from the choice of KLFitter jet selection mode used to admit jets into the fit, see Figures 8.1 and 8.14. By excluding some low p_T light jets, it is possible the best permutation is never obtained. In this study, b-tagging information is used via the choice of kBtagPriority mode. The effect of this choice is analyzed in Section 8.2.3.

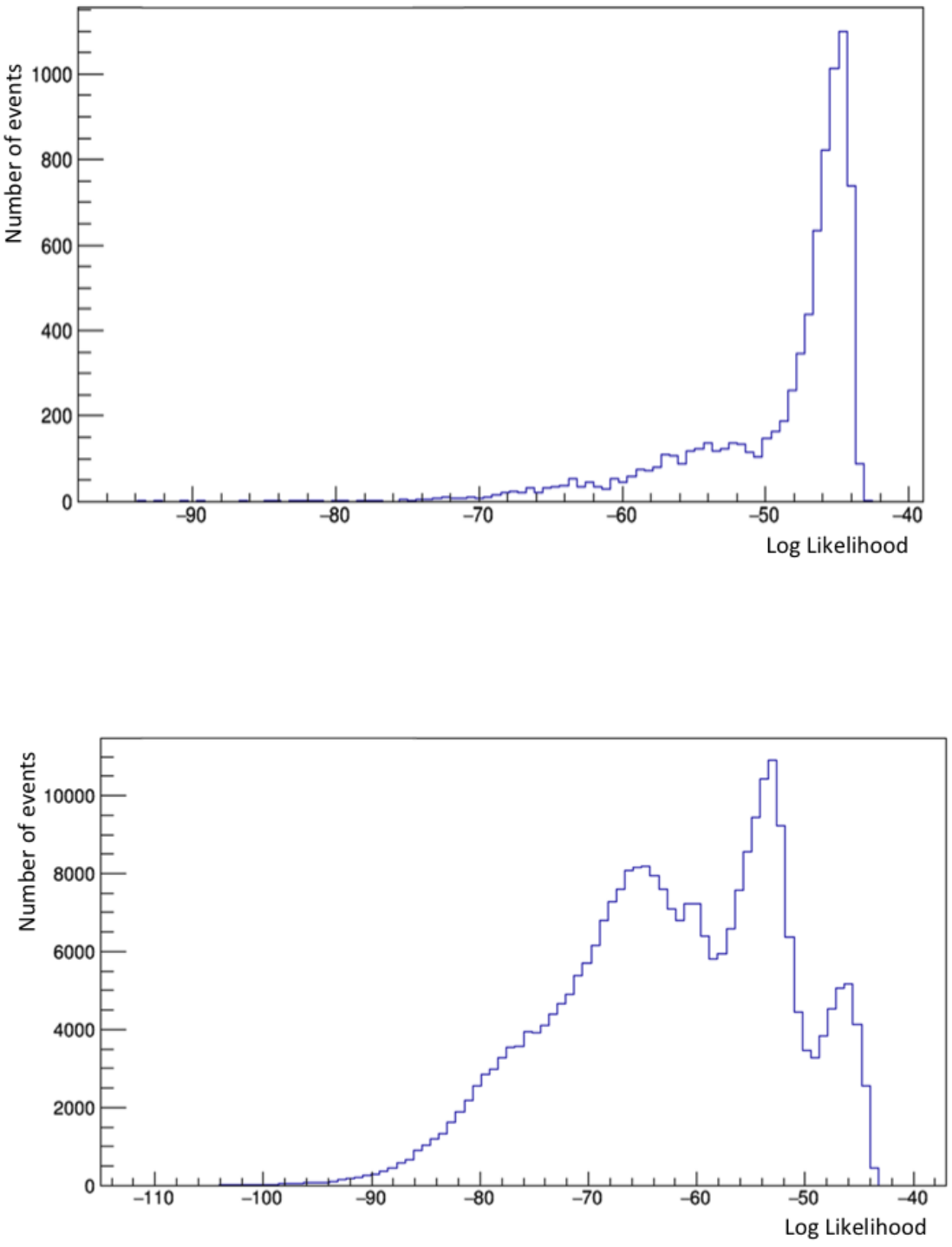


Figure 7.3: Log likelihood for the best permutation (upper plot) and all 12 permutations (lower plot). The best permutation from the upper plot corresponds to the last right-hand peak in the 12 permutation plot.

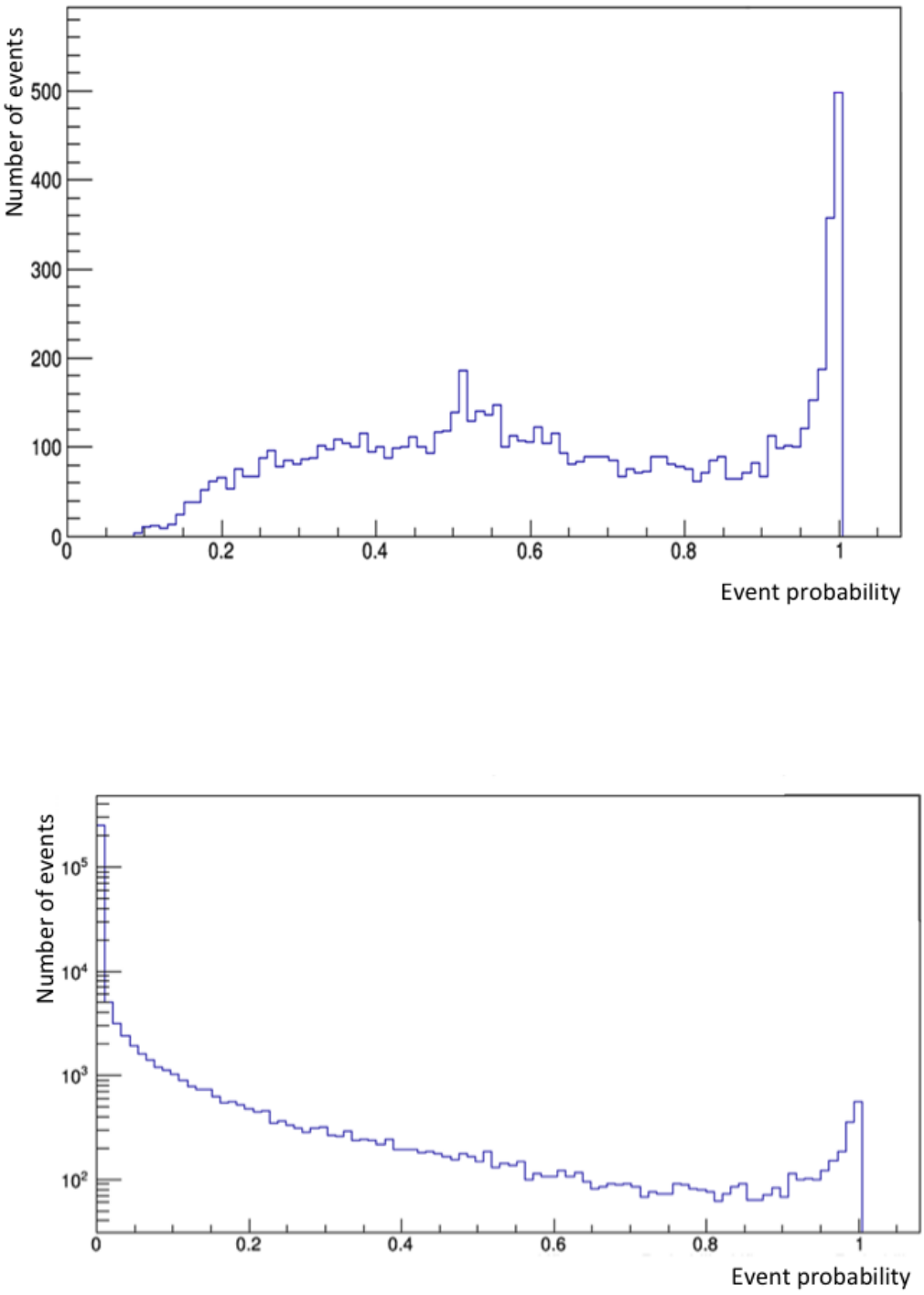


Figure 7.4: Event probability for the best permutation (upper plot) and all 12 permutations (lower plot).

7.4.2 Transfer Functions

Transfer functions model the detector response as well as acceptance and efficiency effects. The transfer functions are derived from the the Next-to-Leading Order (NLO) Monte Carlo program MC@NLO $t\bar{t}$ signal at an input mass of $m_{top} = 172.5$ GeV [134], based on reconstructed objects that are matched to their predicted generator quarks and leptons using a separation of $\Delta R \leq 0.4$ between a quark and the corresponding jet.

The likelihood of measuring a certain value given a truth value of E or p_T for any of the objects is:

$$\mathcal{L}_W = W((\tilde{E}, \tilde{p}_T)_{\text{reco}} | (E, p_T)_{\text{truth}}). \tag{7.10}$$

KLFitter provides several such energy transfer functions for light jets, b-jets, electrons and transverse momentum transfer functions for muons and MET. The momentum and energy distributions of these particles and jets are asymmetric except for the MET distribution. The transfer function for b-quark jet energies has a significantly longer tail than the one for light-quark jet and lepton energies, since the b-quark jets frequently contain a neutrino and/or a muon from the semileptonic decay of a B or D hadron within the jet. In order to be able to take the asymmetric tails into account, the transfer functions are fitted using the sum of two independent Gaussian distributions (double Gaussian) which includes in total two means μ_1, μ_2 and two standard deviations σ_1, σ_2 and an amplitude, encoded as follows:

$$W(\Delta E) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \left[e^{\frac{-(\Delta E - p_1)^2}{2p_2^2}} + e^{\frac{-(\Delta E - p_4)^2}{2p_5^2}} \right] \tag{7.11}$$

with $\Delta E = (E_{\text{true}} - E_{\text{reco}})/E_{\text{true}}^2$, and parameters p_i depend them selves on E_{true}. As an example, the parameterization of b and light jets is as follows:

$$p_1 = a_0 + a_1/E_{\text{true}} \tag{7.12}$$

$$p_2 = a_2 + a_3/\sqrt{E_{\text{true}}} \tag{7.13}$$

$$p_3 = a_4 + a_5/E_{\text{true}} \tag{7.14}$$

$$p_4 = a_6 + a_7/\sqrt{E_{\text{true}}} \tag{7.15}$$

$$p_5 = a_8 + a_9 \cdot E_{\text{true}} \tag{7.16}$$

where the parameters a_i are obtained from the fits. Similar parameterization exist for the other objects.

The transfer function parameterization is chosen differently for separate η-regions except for the one for missing transverse momentum. For electrons, b-jets and

light-jets, these regions at $\sqrt{s}$ =8 TeV are:

$$|\eta| < 0.8 \tag{7.17}$$
$$0.8 < |\eta| < 1.37 \tag{7.18}$$
$$1.37 < |\eta| < 1.52 \tag{7.19}$$
$$1.52 < |\eta| < 2.5 \tag{7.20}$$

For muons, are chosen three η-regions are chosen:

$$|\eta| < 1.11 \tag{7.21}$$
$$1.11 < |\eta| < 1.25 \tag{7.22}$$
$$1.25 < |\eta| < 2.5 \tag{7.23}$$

7.4.3 Breit-Wigner function

The factor which denotes the probability for a particle to have a reconstructed mass $\tilde{m}$ in the likelihood function, given the SM width and pole mass predictions, is considered as the mass constraint. The natural width of unstable particles may vary over the relativistic BreitWigner resonance shape 7.24 which is characterized by a sharper peak and flatter tails of the curve than a comparable Gaussian. Longer-lived particles have smaller width and hence sharper peaks with a lifetime which is correlated to the inverse of the uncertainty (i.e. the width) in its rest mass. Therefore, the likelihood exploits the values of m_W and Γ_W to constrain the reconstructed leptonic, $m(l\nu)$, and hadronic, $m(q1q2)$ W-boson masses using Breit-Wigner distributions. Similarly, the reconstructed leptonic, $m(l\nu b_{lep})$, and hadronic, $m(q1q2b_{had})$, top-quark masses are fitted to the SM top mass and width Γ. The Higgs boson reconstructed mass $m(b1Higgsb2Higgs)$ is unconstrained by the fit, but the reconstructed mass enters into the likelihood $\mathcal{L}$ eq.(7.4) through the BW distribution for SM Higgs of mass m_{Higgs} and width Γ_{Higgs}. The relativistic Breit-Wigner distribution is defined as:

$$BW(\tilde{m}|m,\Gamma) = \frac{1}{(\tilde{m}^2 - m^2)^2 + m^2\Gamma^2} \tag{7.24}$$

Where m and Γ are the nominal mass and width of the particle and $\tilde{m}$ is the mass of the reconstructed particle [135].

7.5 Summary

This chapter described the operation of the KLFitter package; including the likelihood function for the $t\bar{t}H(H \to bb)$ channel. The KLFitter modes of operation using b-tagged jets or simply p_T-ordered jets were detailed. The calculation of the 'event probability' for each permutation was also shown. The ranking of these permutations leads to a best estimate for the choice of final-state particle assignments. An improvement to the process was introduced, which allows more jets to be considered in the permutations. Such that, the efficiency of the fitter improves slightly.

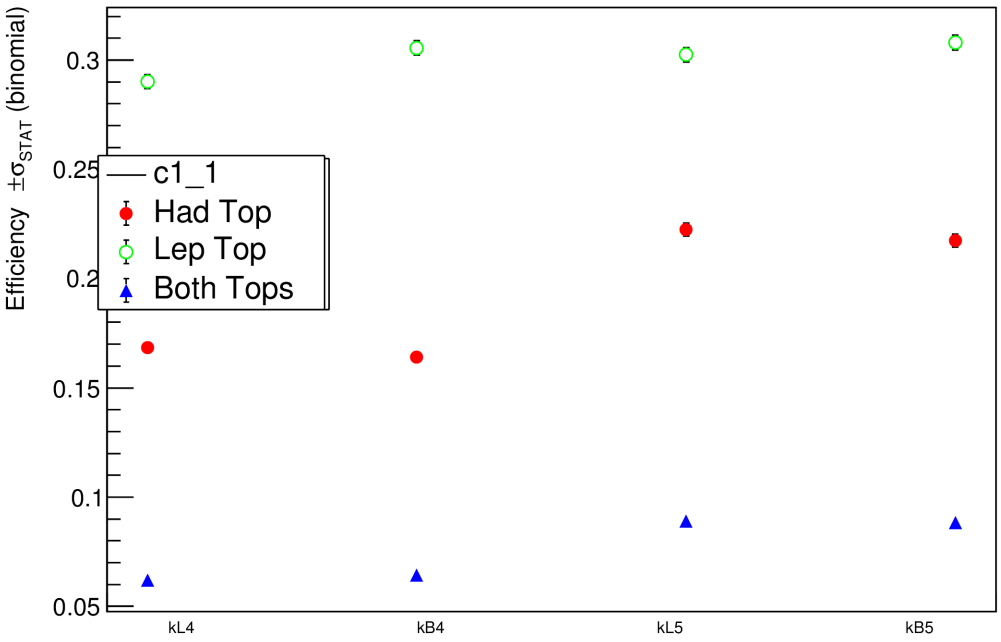
Chapter 8

Performance of the Kinematic Fit

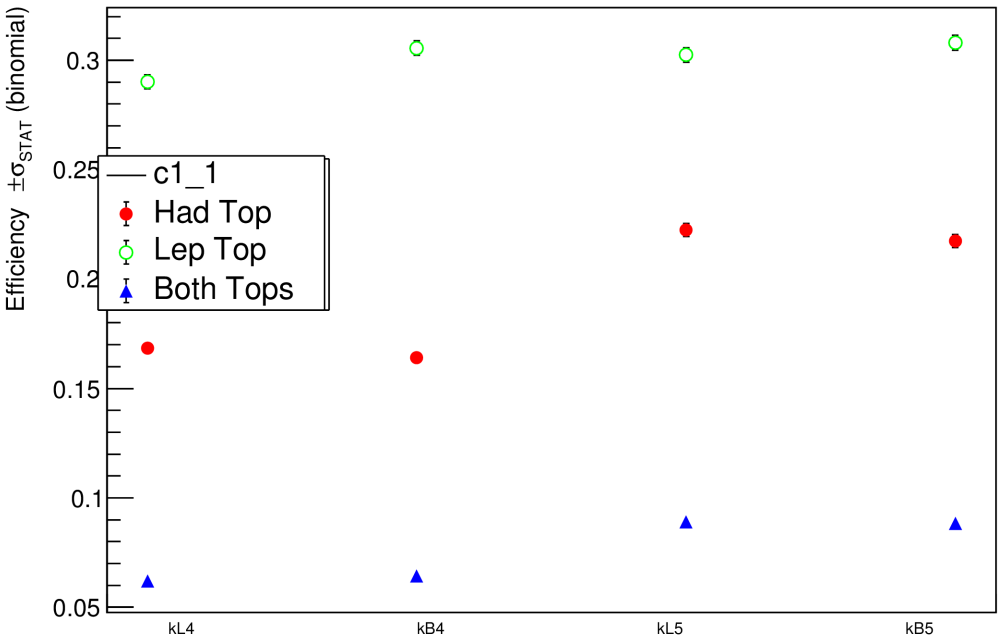
This study uses the TTHbbLeptonic package to run KLFitter and perform event reconstruction. The analysis of the output is performed using custom code which is written from scratch. Starting in Section 8.1 using KLFitter to reconstruct $t\bar{t}$ events only, the objective is to modify the inputs and configuration parameters and understand the output. The analysis code performs a matching of reconstructed jets to truth partons in the event, in order to assess the efficiency of the fitter in various KLFitter running modes, starting with the kL4, kL5, kb4 and kb5, describing in Table 7.1. Then, the study is extended to $t\bar{t}H$ events: initially using the original KLFitter likelihood to perform the reconstruction of the two top quarks, and assessing the performance. Eventually the KLFitter likelihood is extended to consider the whole $t\bar{t}H$ system by using the remaining b-tagged jets to reconstruct the Higgs boson. This step extends the KLFitter likelihood to KLFitterLH, and is covered in Section 8.2. The KLFitterLH package implements the four new jet selection modes; kL6, kB6, kB7 and kB8, see Table 7.1. The goal is to measure the reconstruction efficiency of the Higgs boson and to decide upon the most powerful KLFitterLH variables that show a good separation between signal and background. These Monte Carlo performance studies are carried out first in a subset of the available events, running locally. In the final step, this was extended to the full statistics available.

8.1 Performance Studies in Four and Five jet selection modes

This study employs the $t\bar{t}$-style four and five jet selection modes of KLFitter described in Section 7.3 (kL4, kL5, kb4 & kb5) to calculate the truth-matching efficiency of the reconstructed hadronic top, leptonic top, and also to examine the efficiency of finding both hadronic and leptonic tops in the same event, see Figure 8.1 (a). The study is extended to study these four basic KLFitter modes for $t\bar{t}H$ events, see Figure 8.1 (b).



(a)



(b)

Figure 8.1: Efficiency for different KLFitter jet selection mode for $t\bar{t}$ events (upper plot) and $t\bar{t}H$ events (lower plot) for the reconstructed hadronic top, leptonic top and both of them together.

Some important observations can be inferred from Figure 8.1. The efficiency of finding the Leptonic Top is higher than the hadronic top in $t\bar{t}$ and $t\bar{t}H$. For the leptonic top, all the modes almost have the same efficiency in either the $t\bar{t}$ and $t\bar{t}H$. The errors-bars in $t\bar{t}H$ are much smaller than in $t\bar{t}$ because we have more statistics in $t\bar{t}H$. The efficiency of finding both tops in $t\bar{t}$ is higher than $t\bar{t}H$ and that indicates that the method must be improved if we are to increase the efficiency in $t\bar{t}H$. It is likely that the extra b-tagged jets from the Higgs boson are responsible for this affecting the hadronic top most strongly. The presence of a Higgs boson has a strong effect on the hadronic top, which, in turn affect the efficiency of finding both tops together.

Figure 8.2 shows the peak around zero for the ΔR between reconstructed and truth hadronic top in the $t\bar{t}$ events. However, most of the events have $\Delta R > 0.3$. The effect of the detector is studied through the energy resolution function of the hadronic top ($\frac{E_{\text{reco}} - E_{\text{truth}}}{E_{\text{truth}}}$). Figure 8.3 shows that the resolution peaks at zero: this implies that the overall reconstructed top versus truth is well calibrated. However, the resolution is not symmetric, there is a resolution tail due to the loss of energy from the hadronic top which escapes detection by emission of unmeasured particles.

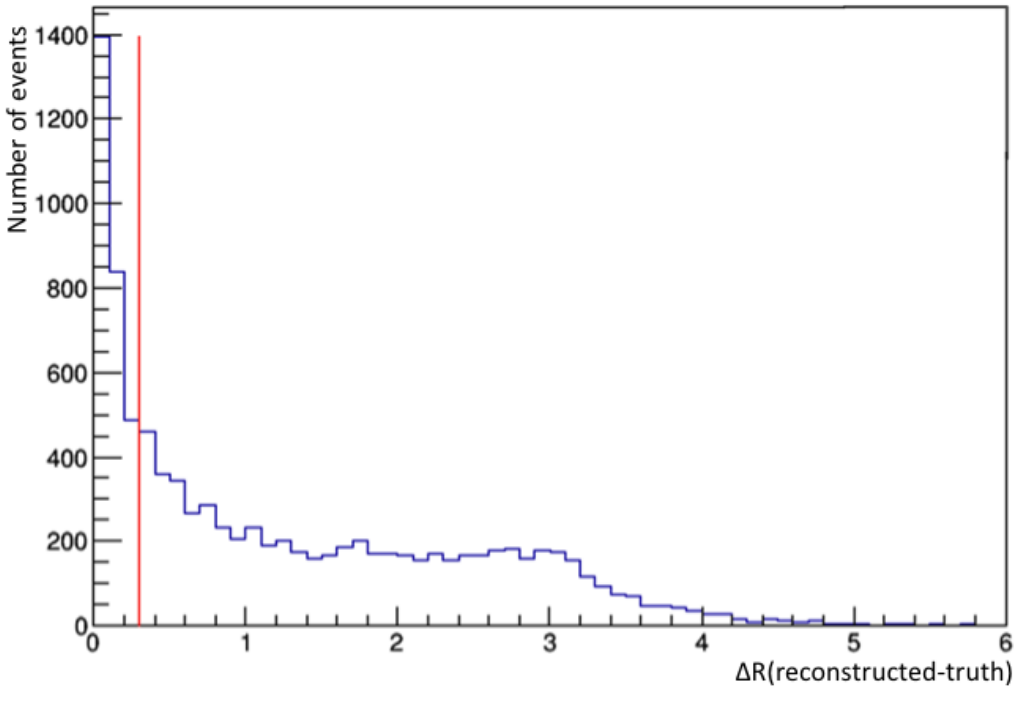


Figure 8.2: The ΔR for the reconstructed hadronic top in $t\bar{t}$, using kBtagPriority-FiveJets mode. The red line indicates the chosen ΔR cut value which is 0.3.

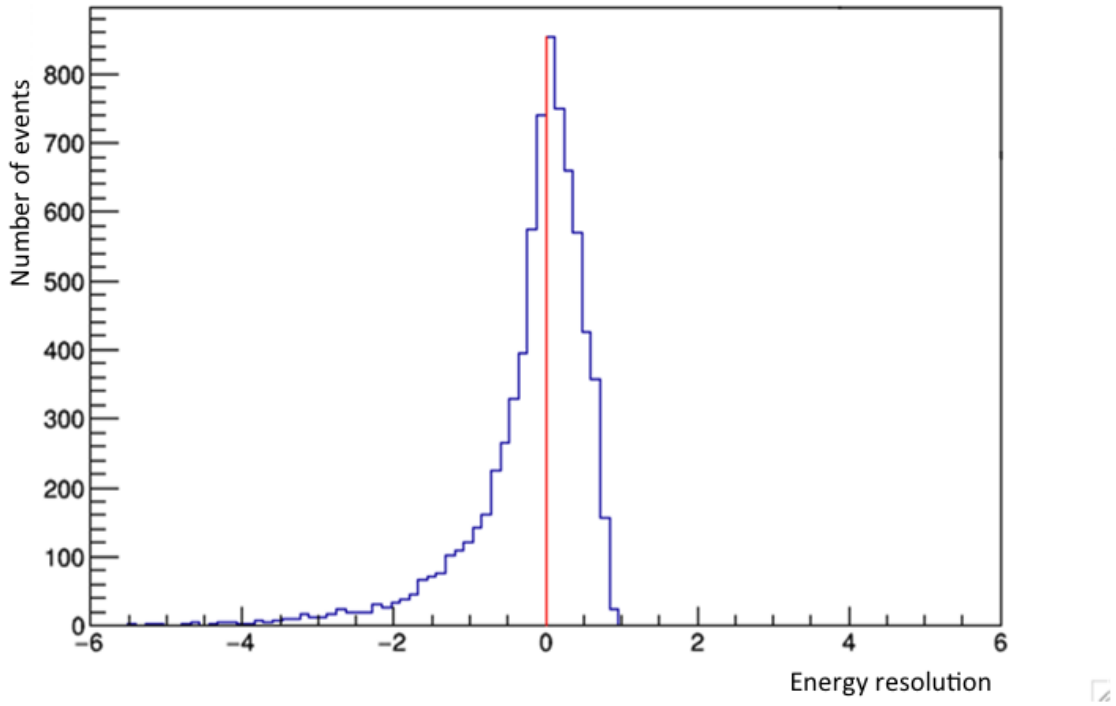


Figure 8.3: The asymmetric energy resolution for the reconstructed hadronic top in $t\bar{t}$ events due to the neutrinos escaping.

8.2 Performance Studies in Six, Seven, and Eight jet selection modes

Section 8.2.1 begins by extending the study to six jets in KLFitter producing 20 default variables as output, see Table 8.1. This stage runs on both $t\bar{t}$+jets and $t\bar{t}H$ events, using the new KLFitter likelihood for six jets system KLFitterLH.

Section 8.2.2 uses increased statistics and trims the output variables from KLFitterLH.

In Section 8.2.3, the study adds extra samples, namely $t\bar{t}H$ dileptonic, and again increases the statistics to the whole available simulated samples. In addition, KLFitterLH is extended to use two new selection modes namely, kB7 and kB8 modes.

8.2.1 Study Stage I

In this Section, the description of the extension of the KLFitter likelihood function performance is presented. The KLFitter jet selection mode is kB6 with the number of events equal to 1558 for the semileptonic $t\bar{t}H$ sample and equal to 529 for the $t\bar{t}$ sample in the (6j,2b) inclusive region. The study commences by looking at the performance of KLFitterLH for 20 output KLFitterLH variables in terms of the separation ability between signal and background, see Table 8.1. The KLFitterLH performance in separating signal $t\bar{t}H$ from background $t\bar{t}$+jets begins to be seen in Figure B.2a which shows reconstructed Higgs mass (H_m^{klf}). The low mass states reconstructed in $t\bar{t}$ events are significantly reduced in the signal sample. The Higgs mass is slightly higher for the signal sample because the invariant mass of the two b-jets from the Higgs decay is, in general, higher than the fake mass in the $t\bar{t}$+jets background (particularly in $t\bar{t}$+cc and $t\bar{t}$+light). It is possible that the $t\bar{t}$+jets sample has a lower Higgs boson mass because the KLFitterLH demands extra jets in the selection and those jets could be ISR or FSR gluons which lower the reconstructed Higgs in the $t\bar{t}$+jets sample. For the $t\bar{t}H$ sample, reconstructed Higgs boson are truth-matched ($\Delta R \leq 0.5$), see Figure 8.15, resulting in 377 matches and 1181 failed matches. Figures B.2b-B.2e also shows the other Higgs variables output from KLFitterLH, namely reconstructed Higgs boson p_T ($H_{p_T}^{klf}$), H_E^{klf}, H_η^{klf} and H_Φ^{klf}. The noteworthy feature is in the Higgs p_T plot where the signal sample has a slightly higher p_T distribution. Figures B.3a-B.3e shows that the $t\bar{t}$ system's mass ($t\bar{t}_m^{klf}$), $t\bar{t}_{p_T}^{klf}$ and $t\bar{t}_E^{klf}$ all have a slightly higher mean in the signal sample $t\bar{t}H$ compared to the $t\bar{t}$+jets background. This is reasonable because of the extra p_T and E coming from the Higgs. For the two transverse momenta of the hadronic top (thad) and leptonic top (tlep), it clear in Figure B.4a and B.5a that there is some separation between signal and background. The thad and tlep energies in Figures B.4b and B.5b, provide very slightly separation between signal and background, but this is not a significant effect and may possibly be accounted for by ISR and FSR gluon jets. As expected, the reconstructed top quarks have symmetric and flat distributions in η and ϕ as shown in Figures B.4d, B.4c, B.5d and B.5c. Finally, the output values of klftter_eventProbability and klfitter_logLikelihood are shown in Figures B.1a and B.1b. It can be seen that the event-probability has a slightly higher mean for the signal sample and the highest bin (klftter_eventProbability$\simeq$1) has the highest proportion of signal to background. B.

Variable	Description
m_H^{KL}	Higgs boson mass
$p_{T,H}^{KL}$	Higgs boson p_T
η_H^{KL}	Higgs boson η
E_H^{KL}	Higgs boson energy
ϕ_H^{KL}	Higgs boson ϕ
$m_{t\bar{t}}^{KL}$	$t\bar{t}$ system mass
$p_{T,t\bar{t}}^{KL}$	$t\bar{t}$ system p_T
$\eta_{t\bar{t}}^{KL}$	$t\bar{t}$ system η
$E_{t\bar{t}}^{KL}$	$t\bar{t}$ system energy
$\phi_{t\bar{t}}^{KL}$	$t\bar{t}$ system ϕ
m_{thad}^{KL}	hadronic top mass
$p_{T,thad}^{KL}$	hadronic top p_T
η_{thad}^{KL}	hadronic top η
E_{thad}^{KL}	hadronic top energy
ϕ_{thad}^{KL}	hadronic top ϕ
m_{tlep}^{KL}	leptonic top mass
$p_{T,tlep}^{KL}$	leptonic top p_T
η_{tlep}^{KL}	leptonic top η
E_{tlep}^{KL}	leptonic top energy
ϕ_{tlep}^{KL}	leptonic top ϕ

Table 8.1: Reconstructed variables used by KLFitter for the Higgs boson, $t\bar{t}$ system, hadronic and leptonic tops.

8.2.2 Study stage II

This study is extended by defining a new variable ΔR_{b1b2} (defined in the last line of Table 8.2) and looking more closely at the previous 20 variables through increasing the statistics to 6744 and 4502 for the $t\bar{t}H$ and $t\bar{t}$+jets samples, respectively. In the inclusive (6j,3b) region, the number of events falls to 3607 and 689 for $t\bar{t}H$ and $t\bar{t}$+jets, respectively, notably increasing the signal to background ratio. This study uses only the kB6 selection mode. In this study, two kinds of cuts are implemented on the KLFitterLH Higgs variables. The matching cut $\Delta R_{(\text{Higgs Reco-Higgs Truth})} \leq 0.5$ and another cut which is formed using the new variable ΔR_{b1b2}, see Figures 8.4b and 8.4c, where b1 and b2 are the KLFitter variables resulting from Higgs decay;

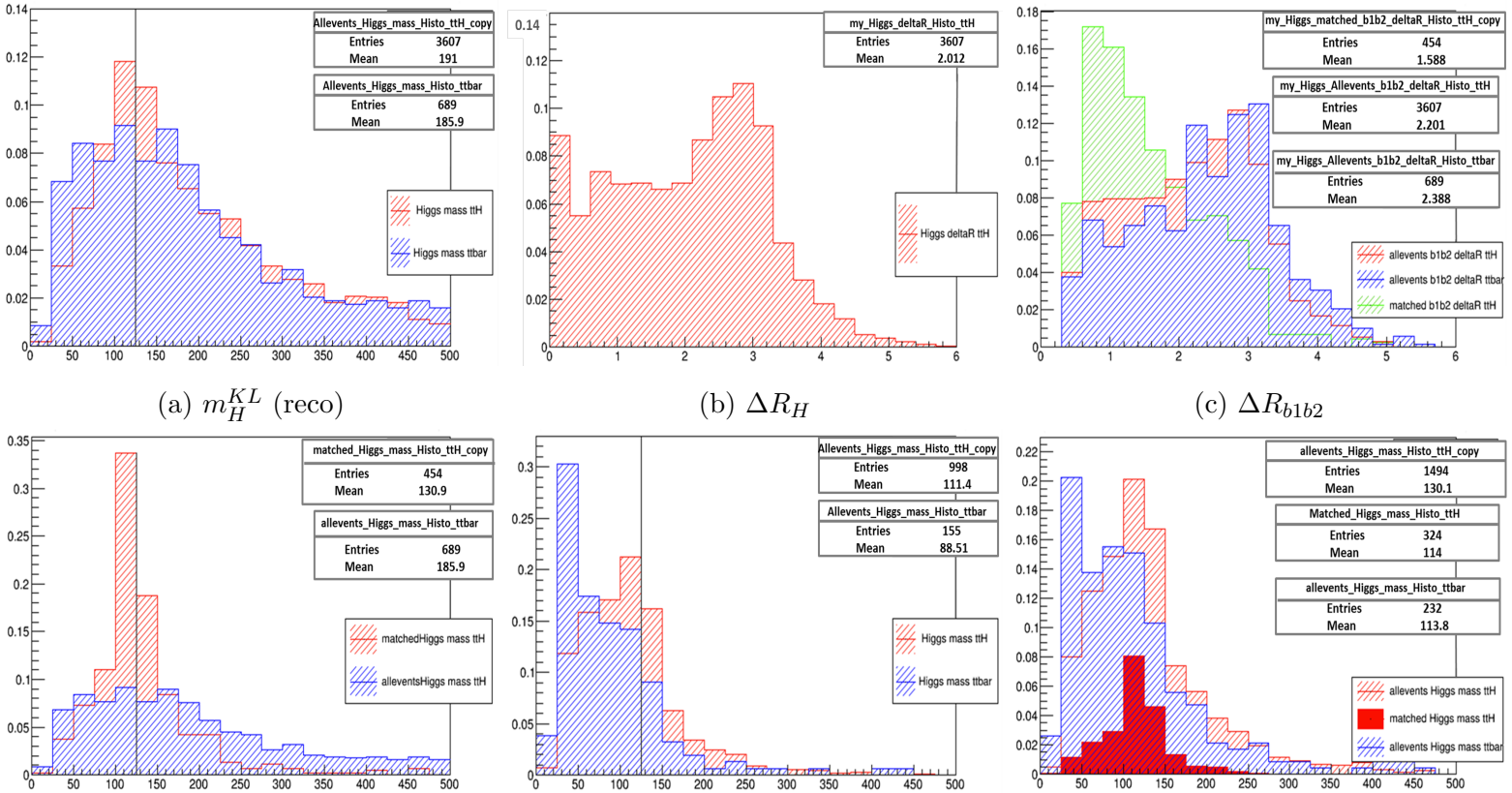
$H \rightarrow b1b2$ and reconstructed by KLFitter. ΔR_{b1b2} provides a way to identify events likely to have a good Higgs truth match without knowing the truth Higgs four vector e.g. in data. To clarify, the Higgs truth matching cut and the ΔR_{b1b2} cut are applied on the Higgs mass variable as shown in Figures 8.4e and 8.4f. The Higgs truth matching cut is applied to $t\bar{t}H$ events only and compared to $t\bar{t}$+jets events. The ΔR_{b1b2} cut is applied to both $t\bar{t}H$ and $t\bar{t}$+jets. The $\Delta R \leq 0.5$ matching cut is chosen with the aid of Figures 8.16a, 8.16b and 8.16c to purify the reconstructed events depending on truth-matching. To identify the subset of $t\bar{t}H$ events, where KLFitterLH produced a truth-matched Higgs, the value of ΔR_{b1b2} is chosen to be $\Delta R_{b1b2} \leq 2$ with the aid of Figure 8.4c. The variable ΔR_{b1b2} is considered useful when confirming that the matched Higgs events comprise the Higgs mass peak around 125 GeV in Figure 8.4f. Without making the truth matching cut, the ΔR_{b1b2} cut can separate signal from background leaving events which have a well-matched Higgs in truth. Therefore, ΔR_{b1b2} is a useful variable for the present study. The previous two kinds of cut are also implemented on the m_H^{KL}, $p_{T,H}^{KL}$, E_H^{KL}, η_H^{KL} and ϕ_H^{KL}. For the two variables $p_{T,H}^{KL}$ and E_H^{KL}, the signal has a higher mean than the $t\bar{t}$+jets fake Higgs distribution when no ΔR cuts are applied, as shown in Figures 8.5a and 8.6a. When the two ΔR cuts are applied individually in Figures 8.5b, 8.5c, 8.6b and 8.6c, both cuts improve the separation of signal vs. background. As for η_H^{KL} distribution, Figures 8.7a, 8.7b and 8.7c, the truth-matched Higgs particles are more centrally at $\eta = 0$ in the detector than either the unmatched or $t\bar{t}$+jets background. The ΔR_{b1b2} cut has a marginal effect in η.

The remaining variables ($p_{T,t\bar{t}}^{KL}$, $E_{t\bar{t}}^{KL}$, $p_{T,thad}^{KL}$, E_{thad}^{KL}, $p_{T,tlep}^{KL}$ and E_{tlep}^{KL}), Figures 8.8, 8.9 and 8.10) are studied in terms of new truth-matching cut. The value value $\Delta R \leq 1$ is chosen with the aid of Figures 8.8c and 8.9c for tlep and thad. This cut value remove more than half of events for each tlep and thad, respectively. This implies that neither of the two top quarks are well reconstructed.

The remaining variables that are not mentioned in this stage but in the previous study are excluded because they not give a clear separation between signal and background. Therefore, this stage of study end-up with 11 variables, given in Table 8.2.

Variable	Description
m_H^{KL}	Higgs boson mass
$p_{T,H}^{KL}$	Higgs boson p_T
η_H^{KL}	Higgs boson η
E_H^{KL}	Higgs boson energy
$p_{T,t\bar{t}}^{KL}$	$t\bar{t}$ system p_T
$E_{t\bar{t}}^{KL}$	$t\bar{t}$ system energy
$p_{T,thad}^{KL}$	hadronic top p_T
E_{thad}^{KL}	hadronic top energy
$p_{T,tlep}^{KL}$	leptonic top p_T
E_{tlep}^{KL}	leptonic top energy
ΔR_{b1b2}	ΔR between the two b-tagged jets which KLFitter assumes result from the Higgs decay

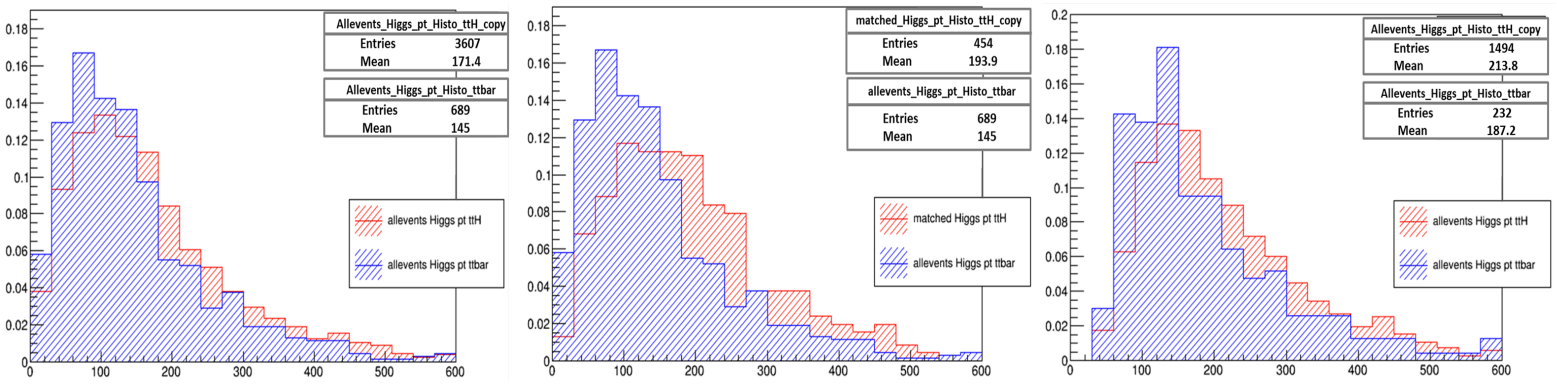
Table 8.2: The subset of reconstructed variables from Table 8.1, plus the addition ΔR_{b1b2} variable used in the second stage of study.



(a) m_H^{KL} (reco) (b) ΔR_H (c) ΔR_{b1b2}

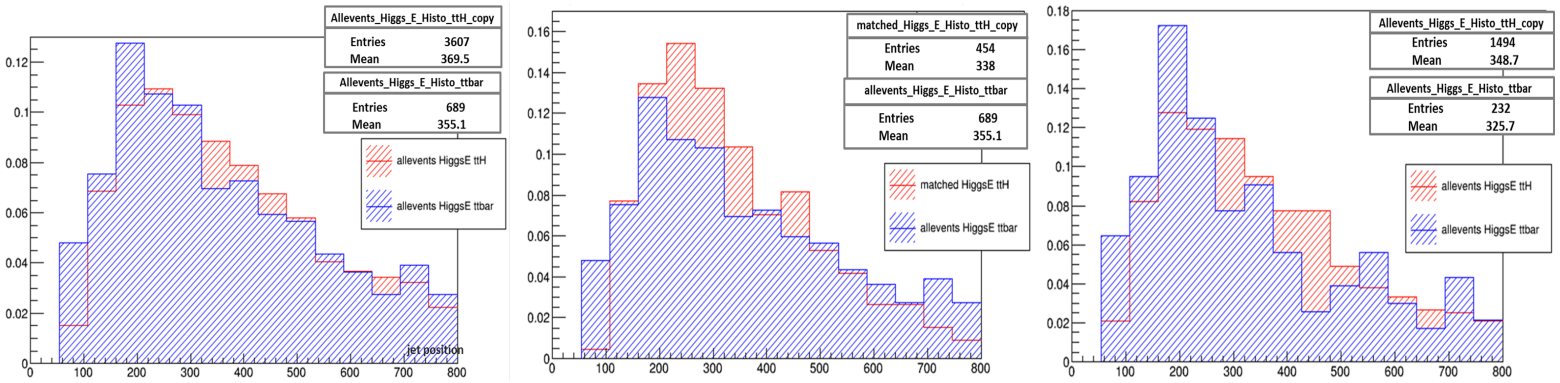
(d) m_H^{KL} (matched-reco) (e) m_H^{KL} and ΔR_{b1b2} cut (f) m_H^{KL} (matched+ΔR_{b1b2} cut)

Figure 8.4: Normalised distributions for the $t\bar{t}H$ and $t\bar{t}$+jets samples, showing: (a) Reconstructed Higgs boson mass [GeV]; (b) $\Delta R_{(\text{Higgs Reco-Higgs Truth})}$; (c) ΔR_{b1b2}. Additionally, shown in green is the distribution for $t\bar{t}H$ events which have a truth-matched ($\Delta R_{(\text{Higgs Reco-Higgs Truth})}$ <0.5) reconstructed Higgs boson; (d) Reconstructed Higgs boson mass for all $t\bar{t}H$ signal and truth-matched (ΔR <0.5) signal; events; (e) Reconstructed Higgs boson mass for $t\bar{t}H$ signal and $t\bar{t}$+jets samples, when a cut on ΔR_{b1b2} ≤2 is applied; (f) Additionaly showing distribution of the truth-matched ($\Delta R_{(\text{Higgs Reco-Higgs Truth})}$ <0.5) reconstructed Higgs boson mass, after the ΔR_{b1b2} ≤2 is applied (solid red). All plots have the cut nBJets≥3 applied.



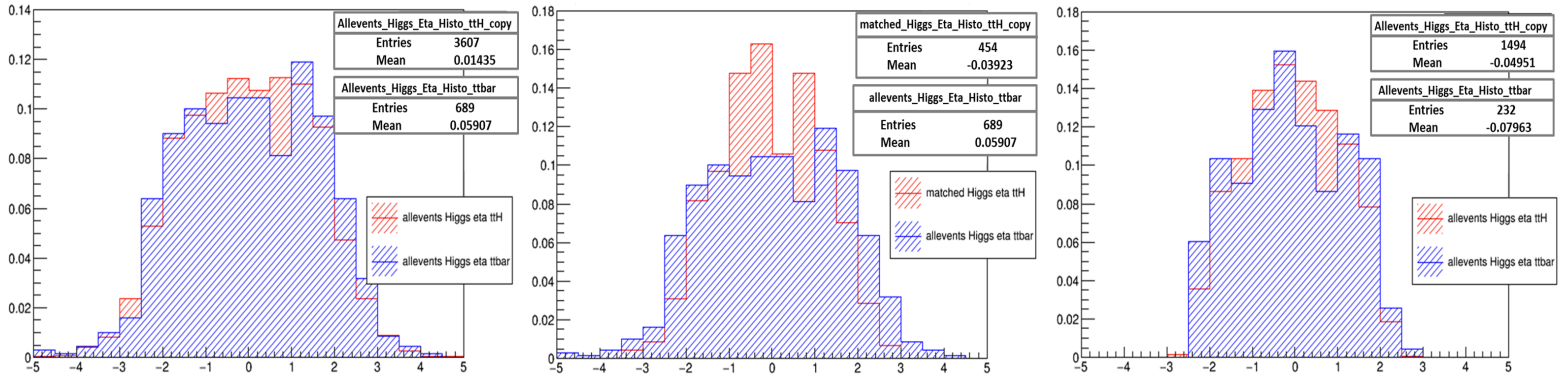
(a) $p_{T,H}^{KL}$ (reco) (b) $p_{T,H}^{KL}$ (matched-reco) (c) $p_{T,H}^{KL}$ (reco-$\Delta R_{b1b2\leq 2}$ cut)

Figure 8.5: Normalised reconstructed Higgs p_T [GeV/c] distributions with a cut on nBJets≥3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples is shown in plot (a). Plot (b) shows the impact on the shape of using only the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) signal sample. Plot (c) shows the impact of applying the ΔR_{b1b2} cut.



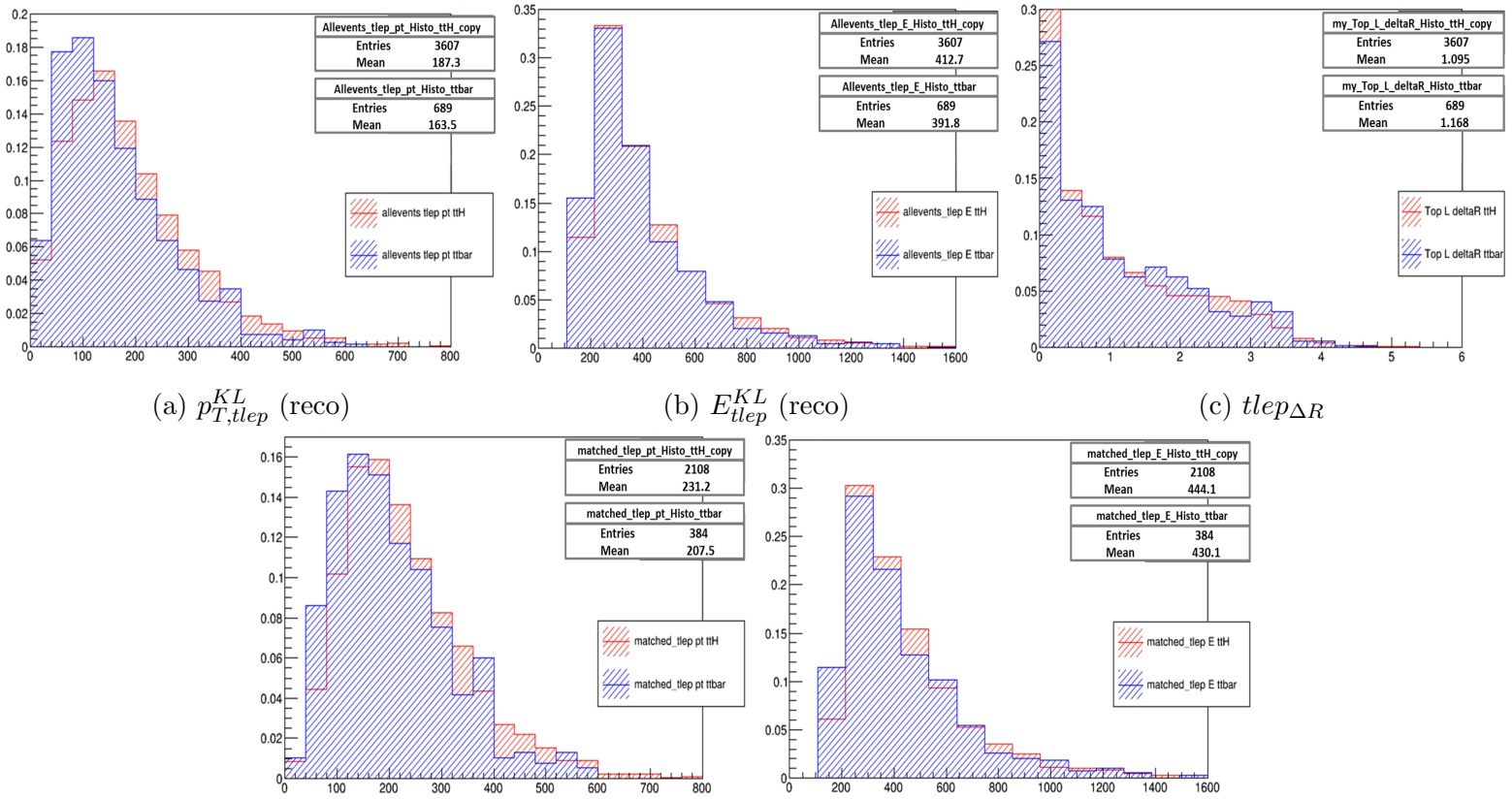
(a) E_H^{KL} (reco) (b) E_H^{KL} (matched-reco) (c) E_H^{KL} (reco-ΔR_{b1b2} cut)

Figure 8.6: Normalised reconstructed Higgs E [GeV] distributions with a cut on nBJets≥3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples. Plot (b) shows the impact on the shape of using only the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) signal sample. Plot (c) shows the impact of applying the $\Delta R_{b1b2} \leq 2$ cut.



(a) η_H^{KL} (reco) (b) η_H^{KL} (matched-reco (c) η_H^{KL} (reco-ΔR_{b1b2} cut)

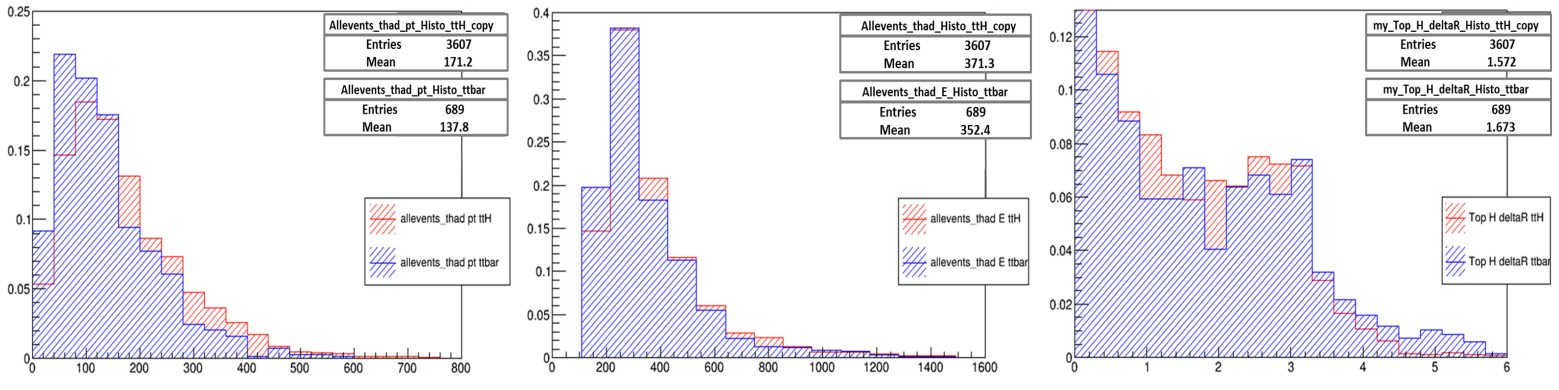
Figure 8.7: Normalised reconstructed Higgs η distributions with a cut on nBJets$\geq$3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples is shown in plot (a). Plot (b) shows the impact on the shape of using only the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) signal sample. Plot (c) shows the impact of applying the $\Delta R_{b1b2} \leq 2$ cut



(a) $p_{T,tlep}^{KL}$ (reco) (b) E_{tlep}^{KL} (reco) (c) $tlep_{\Delta R}$

(d) $p_{T,tlep}^{KL}$ (matched) (e) E_{tlep}^{KL} (matched)

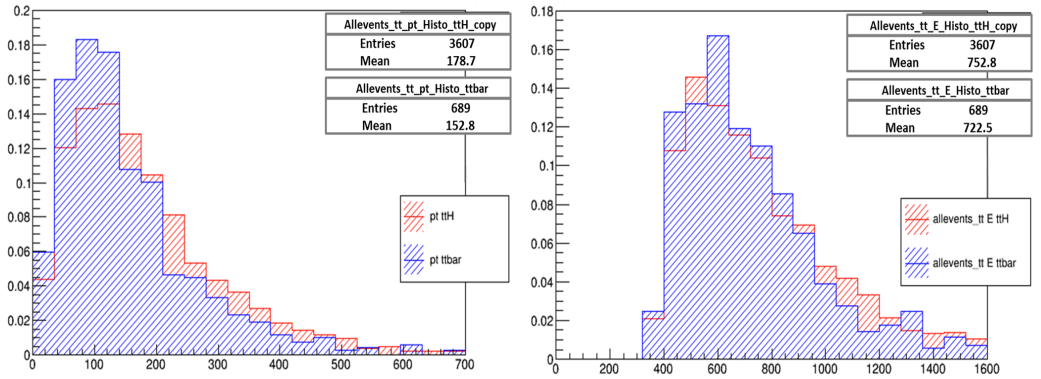
Figure 8.8: Normalised reconstructed tlep p_T [GeV/c] and E [GeV] distributions with a cut on nBJets≥3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples are shown in Figures (a) and (b), respectively. Plots (d) and (e) show the impact on the shape of using the truth-matched ($\Delta R_{\text{truth-reco}} \leq 1.0$) signal and background samples.



(a) $p_{T,thad}^{KL}$ (reco) (b) E_{thad}^{KL} (reco) (c) $thad_{\Delta R}$

(d) $p_{T,thad}^{KL}$ (matched) (e) E_{thad}^{KL} (matched)

Figure 8.9: Normalised reconstructed thad p_T [GeV/c] and E [GeV] distributions with a cut on nBJets≥3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples are shown in figures (a) and (b), respectively. Plots (d) and (e) show the impact on the shape of using the truth-matched ($\Delta R_{\text{truth-reco}} \leq 1.0$) signal and background samples.



(a) $p_{T,t\bar{t}}^{KL}$ (reco) (b) $E_{t\bar{t}}^{KL}$ (reco)

(c) $p_{T,t\bar{t}}^{KL}$ (matched) (d) $E_{t\bar{t}}^{KL}$ (matched)

Figure 8.10: Normalised reconstructed $t\bar{t}$ system p_T [GeV/c] and E [GeV] distributions with a cut on nBJets≥3, for $t\bar{t}H$ signal (red) and $t\bar{t}$+jets (blue) samples are shown in Figures (a) and (b), respectively. Plots (c) and (d) show the impact on the shape of using the truth-matched ($\Delta R_{\text{truth-reco}} \leq 1.0$) signal and background samples.

8.2.3 Study stage III

The plots presented at this stage represent the events in the inclusive (6j,4b) region, using two Monte-Carlo $t\bar{t}H$ and two $t\bar{t}$+jets samples, each having $\sqrt{s}$= 13 TeV. The three KLFitterLH selection modes kB6, kB7 and kb8 are used, and the final reduction of the variables from 11 to 5 is also made at this stage, see Table 8.3. The study commences by comparing the efficiencies (i.e. successfully reconstructing the matched Higgs, tlep, thad, and both top quarks in the signal, $\epsilon_{\text{efficiency}} = \frac{\text{Matched ttH events for each variable}}{\text{All events}}$) of two of the KLFitterLH selection modes, which are kL6 and kB6. The results are displayed in Figure 8.11:

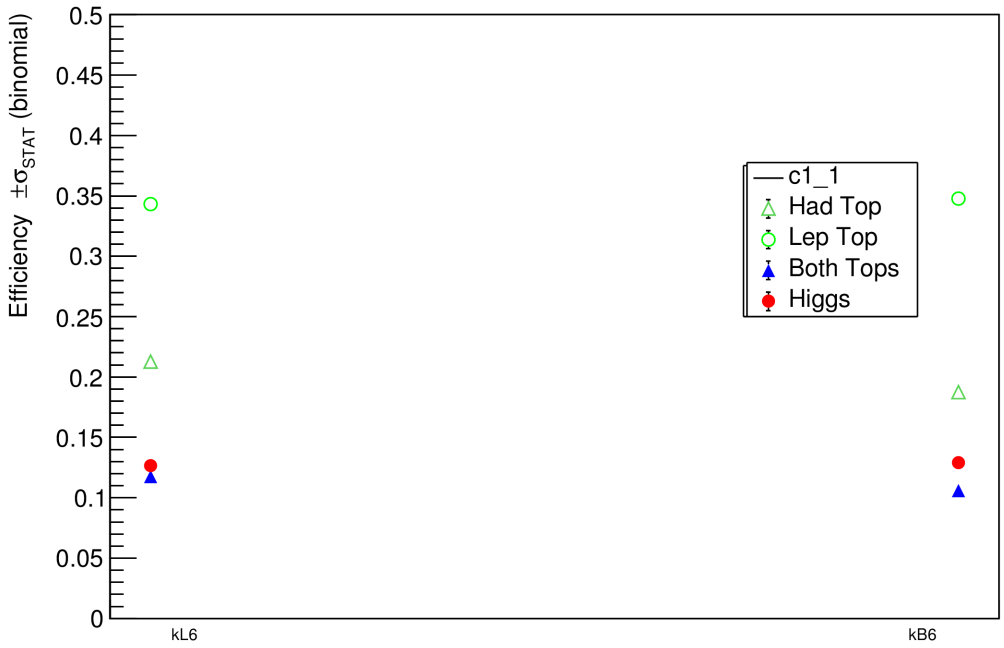


Figure 8.11: Truth matching efficiency for signal events in two KLFitterLH selection modes.

The kL6 and kB6 have almost the same efficiency, but kB6 has a slightly higher efficiency for finding the Higgs. Consequently, the kB mode is chosen in the up coming KLFitterLH analysis. It is found that the efficiency of finding thad is lower in kB6 than kL6. A possible explanation of for this is that in kB6, KLFitterLH has to assign the b-tags first, and if KLFitterLH does not have enough b-tags, then it takes the light jets instead to assign them to the top quarks. Then, other light jets will be wrongly assigned to the W-boson. However, in kL6 mode, the KLFitterLH assigns the jets depending only on the p_T ordering and has a higher chance to select

the correct light jets for the W-boson, improving the thad efficiency. The second important observation from the previous efficiency study in Section 8.2.2 is the efficiency of reconstruction of the matched Higgs in the two cases is still very low and this is also clear from the Higgs ΔR plot Figure 8.4b These two observations raise the following two questions; does kB6 provide the necessary b-tag to make Higgs? do the two b-tagged jets assigned to the Higgs decay come from the hadronic top? To answer these questions, the following study is formed by using the number of failed match events to the truth Higgs boson (1181 events from the first stage of the study in Section 8.2.1). This study is formed by counting the truth matches and failed truth matches ($\Delta R \leq$1.0) for thad, tlep and both thad and tlep together, as summarized in Figure 8.12. The number of events where the thad and tlep are both truth matched in the same event leads to the following important deduction: since most of the events have four b-tags (few of them have 5 b-tags and fewer have 6 b-tags), if two of the b-tags are correctly used to make thad and tlep in 306 events, then the remaining number 875 events do not contain well-reconstructed Higgs. Therefore, if a quarter of events fail the Higgs-match and still make good tops, that implies that the KLFitterLH cannot make a suitable Higgs from the selected 6 jets and it is desirable to include more jets. Accordingly, the study is extended by including an extra slot for the candidate jet by choosing the kB7 mode of KLFitterLH and an extra two slots for the kB8.

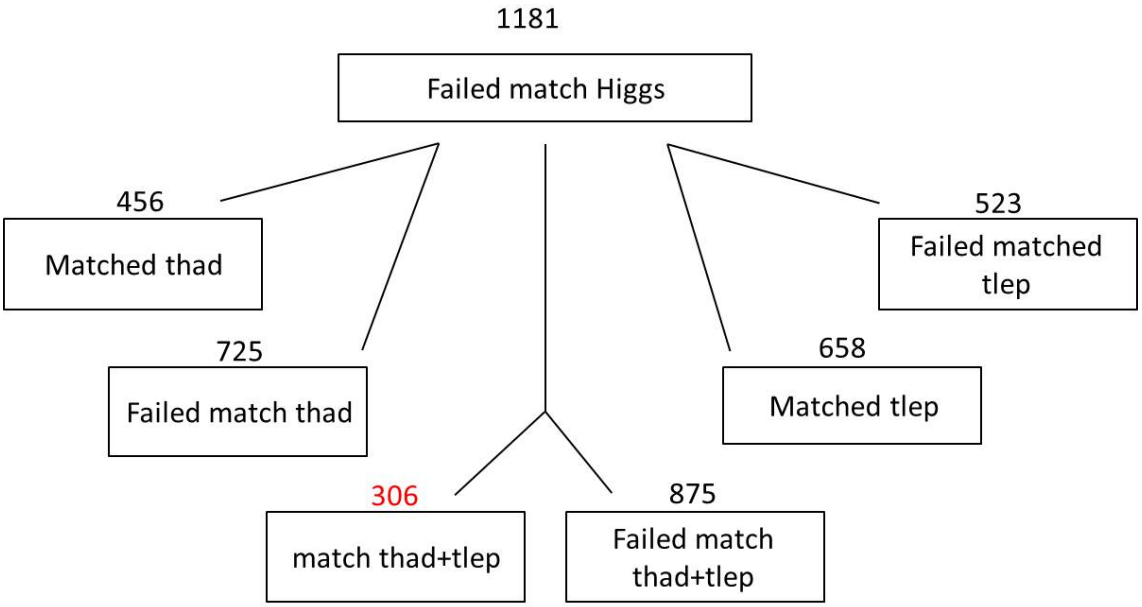


Figure 8.12: Matching efficiencies for the hadronic top (thad), leptonic top (tlep) and both thad and tlep together, for events where the Higgs has failed to match.

At this stage, the KLFitter b-tagging method kVetoNoFitLight is imposed, as dis-

cussed in Section 7.4 which should requests at least 4 b-tagged jets, otherwise the method will fail. The truth-matching efficiencies of Higgs, tlep, thad and both tops together in the signal and background $t\bar{t}$+jets are studied in the three KLFitterLH jet selection modes that are shown in Figures 8.13 and 8.14. In this efficiency study, the efficiency of tlep is higher, as expected, since the electrons and muons are more accurately measured than jets. The Higgs efficiency is around 13% and below thad for the six-jet mode. As more jets are included, the Higgs efficiency increases substantially, even exceeding the efficiency of thad in the 7 and 8 jet selection modes, and in general, all the particles are more efficiency reconstructed as more jets are admitted.

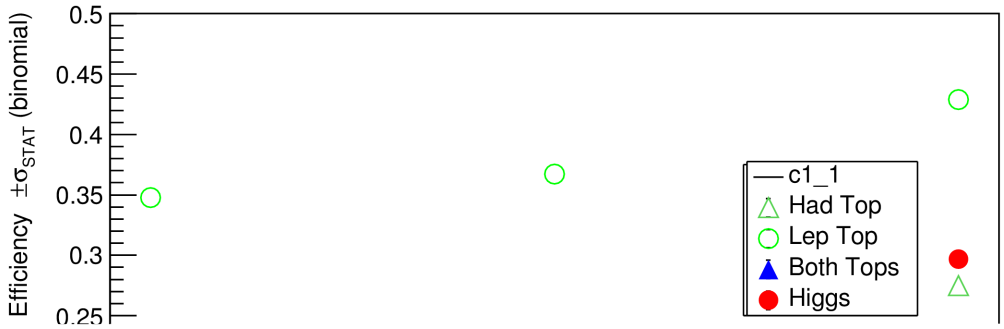


Figure 8.13: Truth matching efficiency for signal events in three KLFitterLH selection modes.

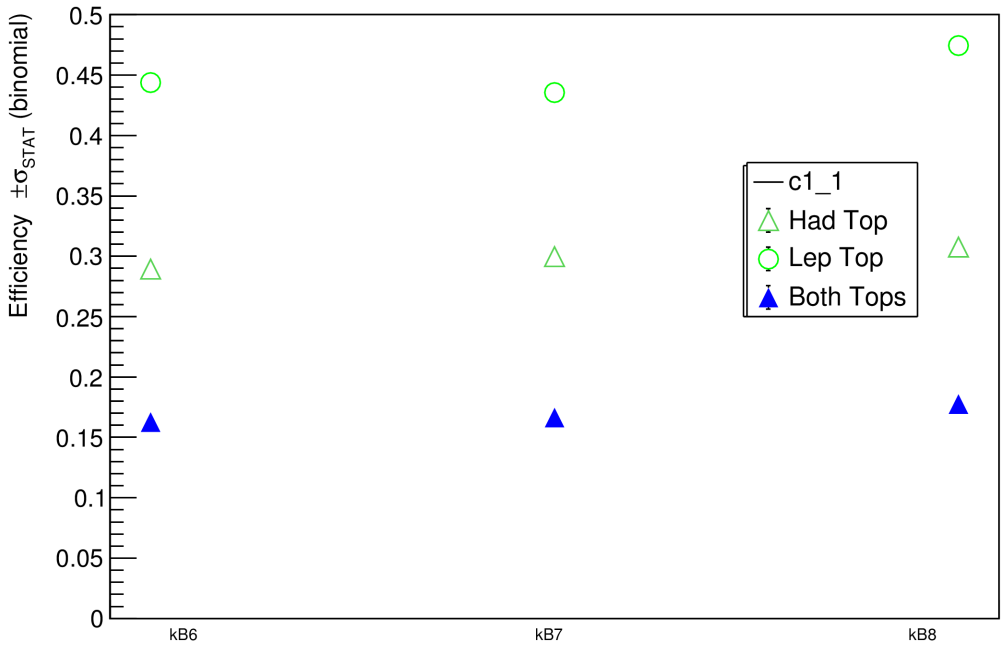


Figure 8.14: Truth matching efficiency for $t\bar{t} + jets$ events in three KLFitterLH selection modes.

To support the efficiency plots (Figure 8.13 and 8.14), the Higgs $\Delta R_{\text{truth-reco}}$ and the new ΔR_{b1b2} variable are analyzed for increased jet multiplicities. The Higgs $\Delta R_{\text{truth-reco}}$ cut becomes more effective at higher jet multiplicities and the $\Delta R_{\text{truth-reco}} \simeq \pi$ peak (back-to-back Higgs) becomes smaller at higher jet multiplicities, see Figures 8.15a, 8.15b and 8.15c. As for the variable ΔR_{b1b2}, the mean is much lower for events with a truth matched Higgs than for events in general. ΔR_{b1b2} is also much lower for truth matched Higgs in signal compared to the fake Higgs in $t\bar{t} + jets$. In addition, for the kB7 and kB8 modes, the $t\bar{t}H$ signal has a lower ΔR_{b1b2} than the $t\bar{t} + jets$ background. This implies that even without truth matching information, the modes can begin to separate signal from background using the ΔR_{b1b2} variable. The useful information that can be gained from this variable is:

- The chance of reconstructing the Higgs correctly is greater when a ΔR_{b1b2} <2 cut is introduced.
- The ΔR_{b1b2} cut at 2, retains 70% of events with a truth-matched Higgs in the all KLFitterLH selection modes.
- It should be a useful variable for the BDT, see Section 9.1.1.

For the Higgs mass m_H^{KL} (Figures 8.17a, 8.17b and 8.17c), the shapes are more promising at kB7 and kB8 jet selection modes because we have a better signal peak, and that is consistent with the efficiency results (Figures 8.13 and 8.20). The matched Higgs plots and Figures 8.18a to 8.18f show that the mass peak for the $t\bar{t}H$ signal is comprised of events where a successful match was made with the truth Higgs boson. However, it is also clear that the background peaks in the same bins. When the ΔR_{b1b2} cut is applied, the Higgs boson mass plots (Figures 8.19a, 8.19b and 8.19c) show that the separation is better than when only a Higgs boson truth-matching cut is applied, but, this cut cannot be applied in data Figure 8.19c shows the Higgs boson reconstructed mass for the kB8 mode which has the most difference in shape because the background is falling smoothly (exponentially) and the signal peaks on this smooth slope. The other interesting observation is that the percentage of Higgs matched events in the mass peak grows between $6 \rightarrow 7 \rightarrow 8$ jets as shown in Figures 8.19a, 8.19b, 8.19c. This matching cannot be used for data, but if the ΔR_{b1b2} variable is used e.g. in the BDT, it will help separation in kB jet modes especially in the kB8 jet mode. Figures 8.18d, 8.18e and 8.18f vs 8.19a, 8.19b and 8.19c show that if the ΔR_{b1b2} <2 cut were to be applied, roughly 70% of the matched events would be retained.

The plots of $p_{T,H}^{KL}$ are shown in Figures 8.21a to 8.21f. The matched events become progressively more separated from the $t\bar{t}$+jets background in the higher KLFitterLH jet selection modes, especially for the kb8 mode. The $p_{T,H}^{KL}$ is therefore a useful variable for the BDT.

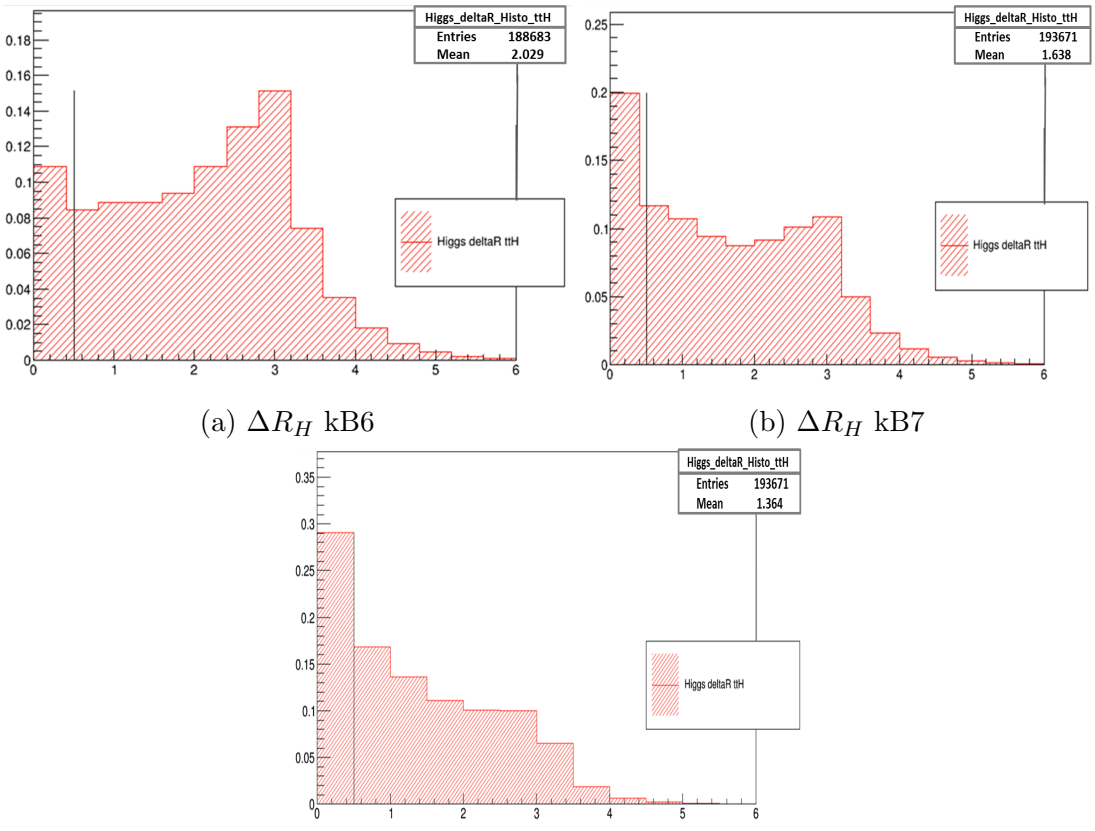
For the E_H^{KL} distributions in Figures 8.22a to 8.22f, the results are somewhat inconsistent across the various KLFitterLH jet modes, with the 7 jet mode exhibiting energies lower for signal than $t\bar{t}$+jets background. The E_H^{KL} could be a useful variable for the BDT.

The other variable showing promise as possible way to separate signal from background is the η_H^{KL}, shown in Figures 8.23a, 8.23b and 8.23c. For truth-match Higgs bosons, η_H^{KL} distribution becomes more central as is shown in Figures 8.23d to 8.24c. In addition, the use of ΔR_{b1b2} information on data gives a clue to the matching.

The other variables, $p_{T,t\bar{t}}^{KL}$, $E_{t\bar{t}}^{KL}$, $p_{T,thad}^{KL}$, E_{thad}^{KL}, $p_{T,tlep}^{KL}$ and E_{tlep}^{KL} are shown in Figures C.1, C.2 and C.3 and are considered to be weak variables in separation ability. The shapes are not affected by changing the kB mode. The shapes of the truth-matched variables are very similar to the unmatched shapes.

Variable	Description
m_H^{KL}	Higgs boson mass
$p_{T,H}^{KL}$	Higgs boson p_T
η_H^{KL}	Higgs boson η
E_H^{KL}	Higgs boson energy
ΔR_{b1b2}	ΔR between two b-tagged jets resulting from Higgs decay and reconstructed by KLFitter

Table 8.3: The subset of reconstructed variables from Table 8.2 used in the third stage of study.

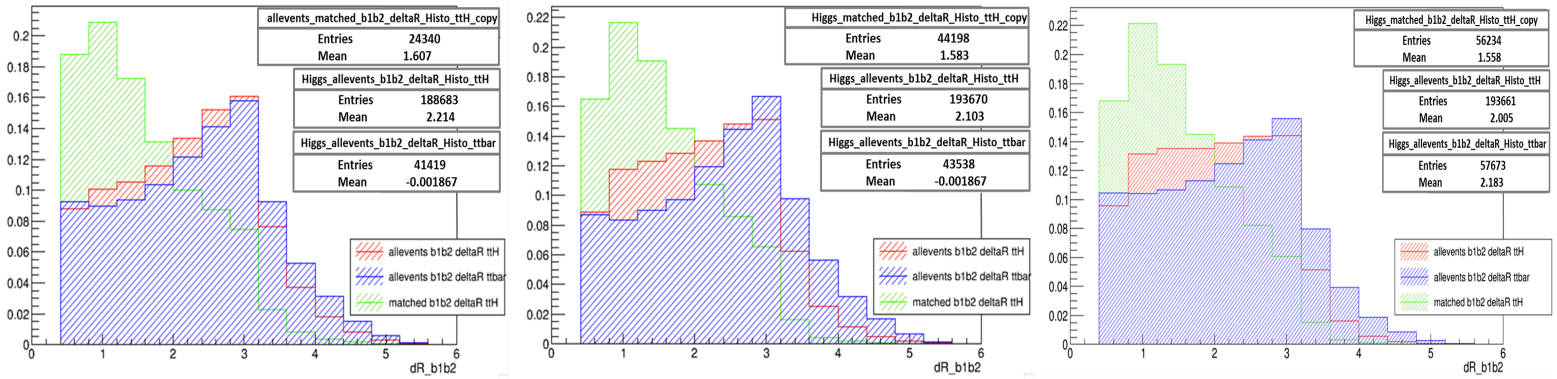


(a) ΔR_H kB6

(b) ΔR_H kB7

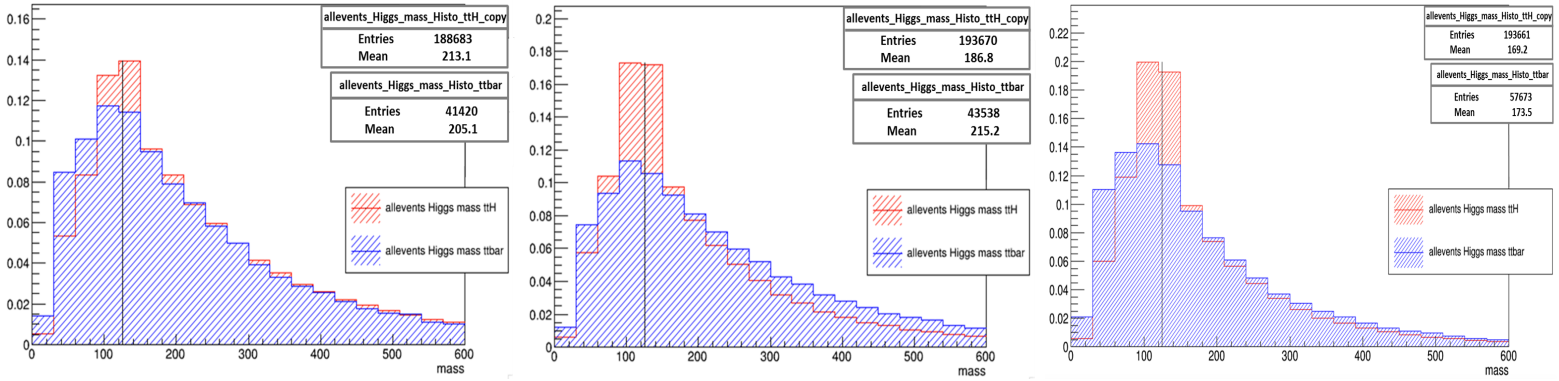
(c) ΔR_H kB8

Figure 8.15: kB6, kB7 and kB8 of Higgs $\Delta R_{\text{truth-maching}}$. The Higgs $\Delta R_{\text{truth-reco}}$ cut becomes more effective at higher jet multiplicities and similarly, the $\Delta R_{\text{truth-reco}} \simeq \pi$ peak (back-to-back Higgs) becomes smaller.



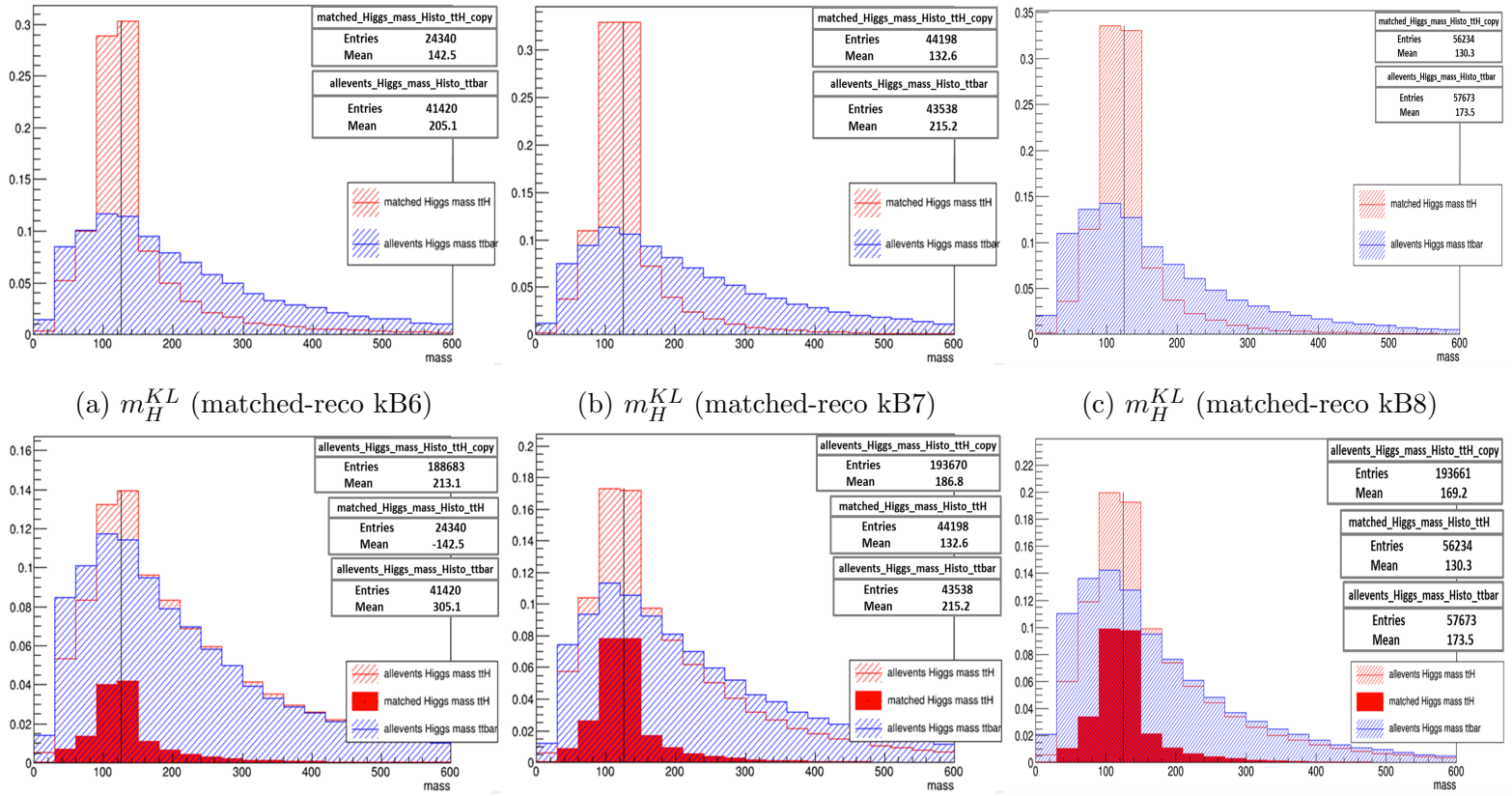
(a) ΔR_{b1b2} kB6 (b) ΔR_{b1b2} kB7 (c) ΔR_{b1b2} kB8

Figure 8.16: Normalized Higgs ΔR_{b1b2} for signal (red) $t\bar{t}$ background (blue) and truth-matched signal events (green). The truth-matched signal distribution differs significantly from the background. In general, the mean Higgs ΔR_{b1b2} reduces as more slots are used by KLFitter in the 6kB, 7kB and 8kB modes respectively.



(a) m_H^{KL} (reco kB6) (b) m_H^{KL} (reco kB7) (c) m_H^{KL} (reco kB8)

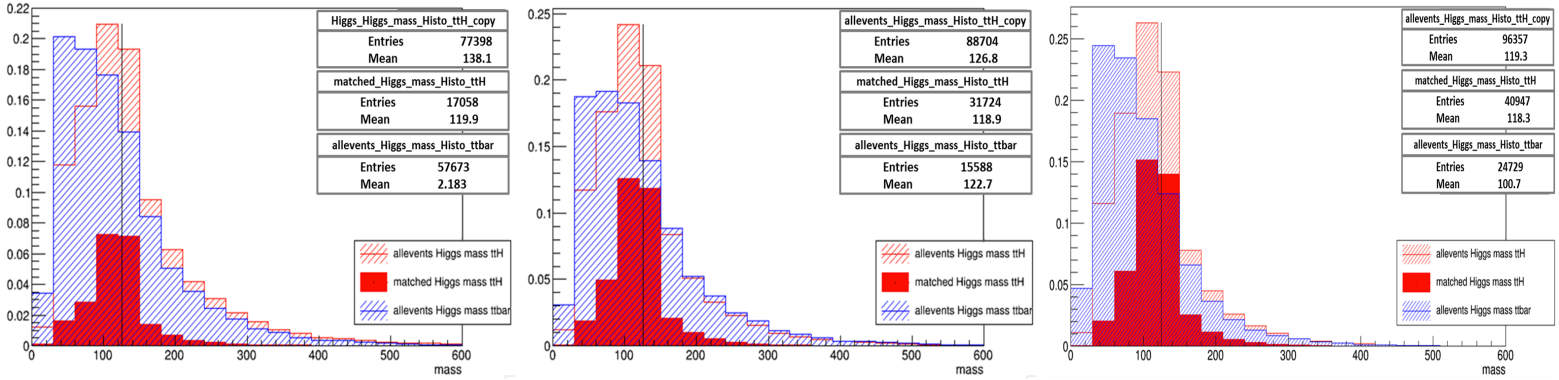
Figure 8.17: The impact of using the kB6, kB7 and kB8 modes of KLFitterLH on the normalized reconstructed Higgs boson mass [GeV] distributions, for $t\bar{t}$+jets (blue) and $t\bar{t}H$ (red) signal samples. The shapes are more promising when using 7 and 8 jets because we have better peak, and that is consistent with the efficiency results from Figures 8.13 and 8.14.



(a) m_H^{KL} (matched-reco kB6) (b) m_H^{KL} (matched-reco kB7) (c) m_H^{KL} (matched-reco kB8)

(d) m_H^{KL} (matched-reco kB6) (e) m_H^{KL} (matched-reco kB7) (f) m_H^{KL} (matched-reco kB8)

Figure 8.18: Upper row: illustration of the effect on the shape of the reconstructed Higgs boson mass [GeV] distributions, when applying the kB6, kB7 and kB8 modes of KLFitterLH on the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (red) and the $t\bar{t}$+jets background samples (blue). Lower row: illustration of the effect on the shape of the reconstructed Higgs boson mass distributions, when applying the kBtagPriority6j/7j/8j modes of KLFitterLH on the $t\bar{t}$+jets background (blue), $t\bar{t}H$ signal (red), and the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (solid red).



(a) m_H^{KL} (matched+reco+ΔR_{b1b2} cut, kB6)

(b) m_H^{KL} (matched+reco+ΔR_{b1b2} cut, kB7)

(c) m_H^{KL} (matched+reco+ΔR_{b1b2} cut, kB8)

Figure 8.19: The impact of using the kB6, kB7 and kB8 modes of KLFitterLH on the normalised reconstructed Higgs boson mass [GeV] distributions, when a cut of $\Delta R_{b1b2} \leq 2$ is applied. The distributions are shown for $t\bar{t}$+jets background (blue), $t\bar{t}H$ signal (red) and the truth-matched $t\bar{t}H$ signal ($\Delta R_{\text{truth-reco}} \leq 0.5$), (solid red).

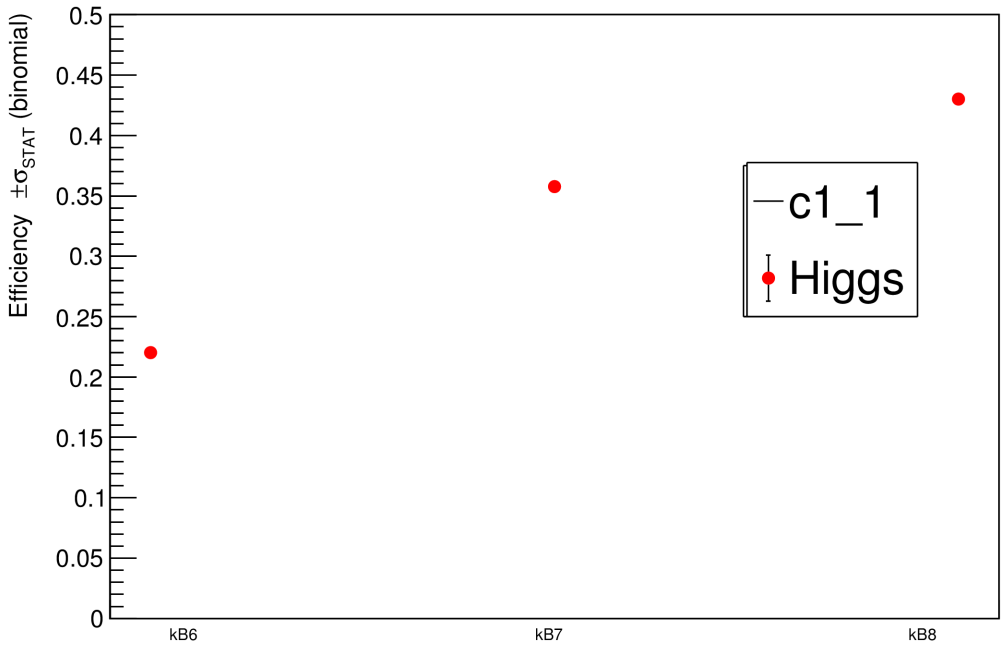
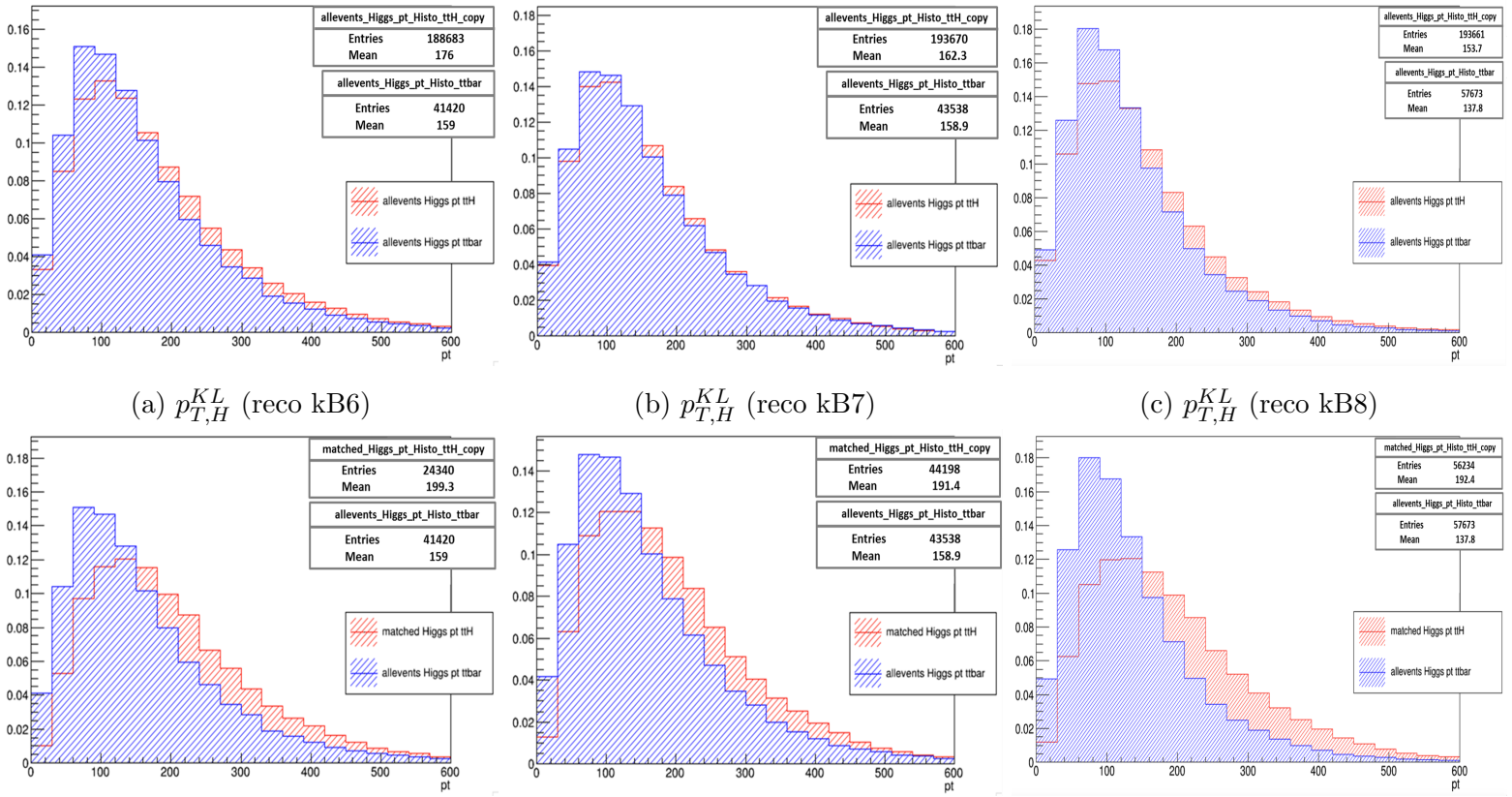


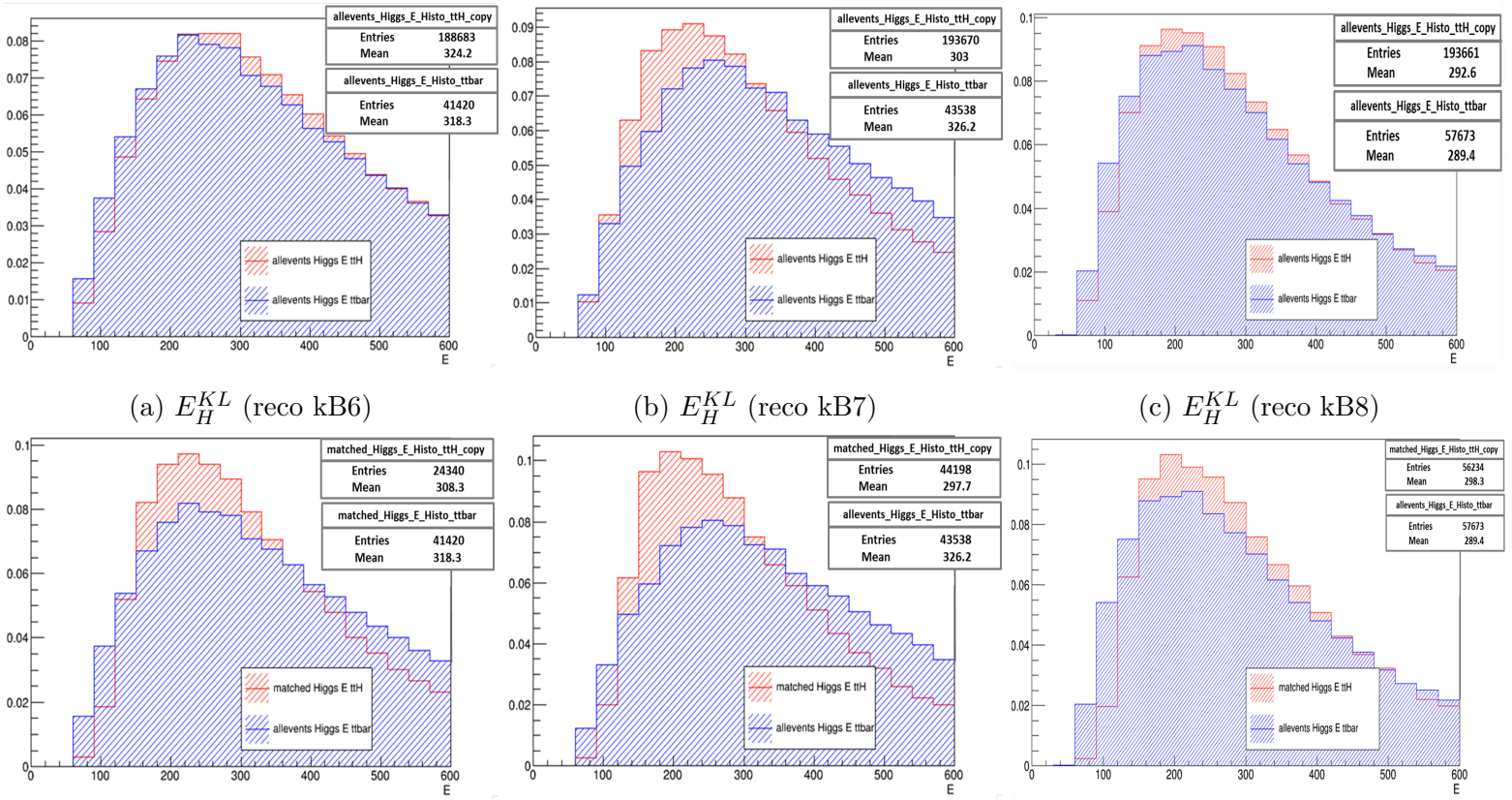
Figure 8.20: Efficiency of kB6, kB7 and kB8 modes when applying Higgs ΔR_{b1b2} cut on matched events.



(a) $p_{T,H}^{KL}$ (reco kB6) (b) $p_{T,H}^{KL}$ (reco kB7) (c) $p_{T,H}^{KL}$ (reco kB8)

(d) $p_{T,H}^{KL}$ (matched-reco kB6) (e) $p_{T,H}^{KL}$ (matched-reco kB7) (f) $p_{T,H}^{KL}$ (matched-reco kB8)

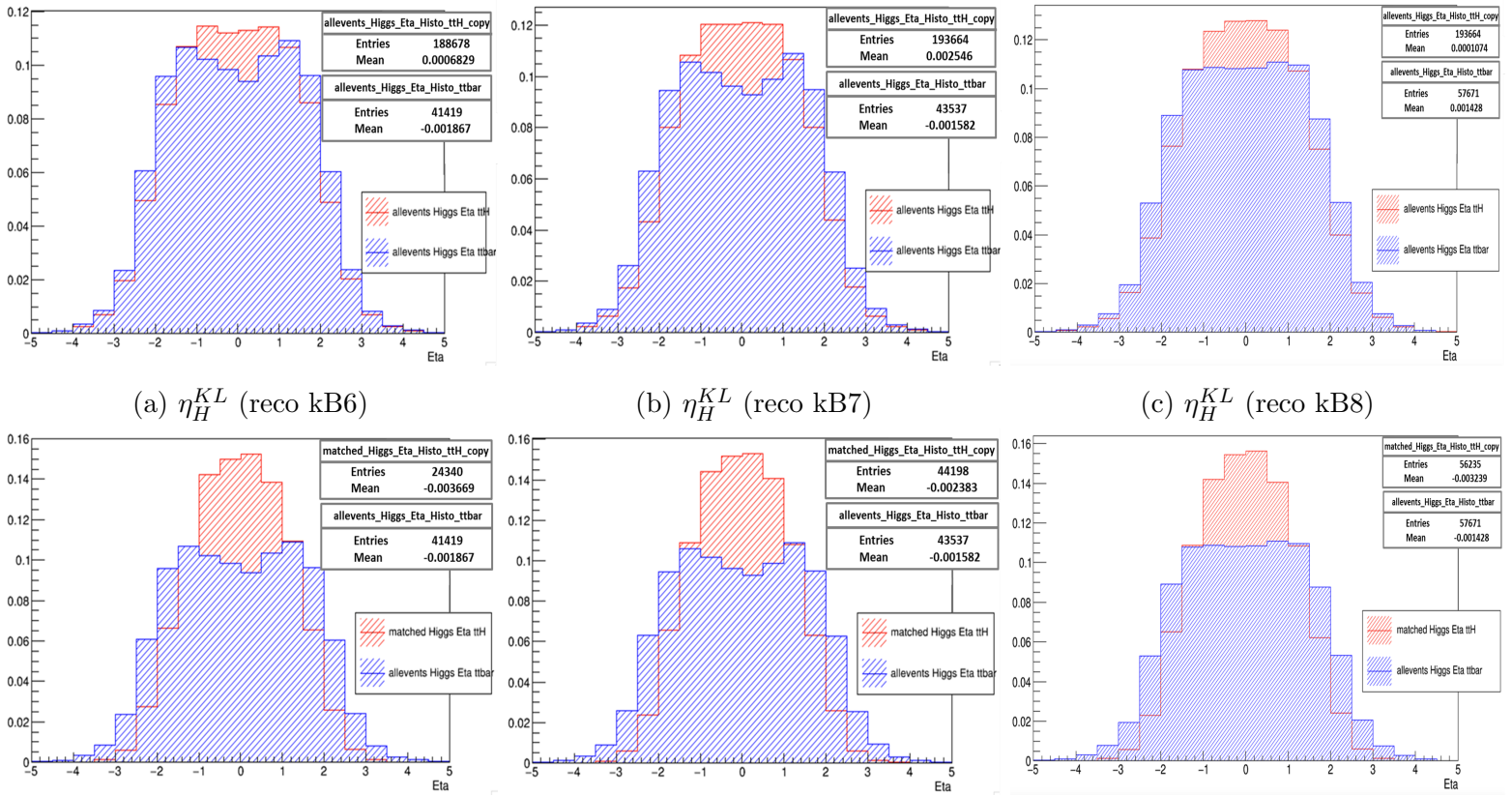
Figure 8.21: Upper row: the impact of using the kBtagPriority6j/7j/8j modes of KLFitterLH on the normalized reconstructed Higgs boson p_T [GeV/c] distributions, for $t\bar{t}$+jets (blue) and $t\bar{t}H$ (red) signal samples. Lower row: illustration of the effect on the shape of the reconstructed Higgs boson p_T distributions, when applying the kB6, kB7 and kB8 modes of KLFitterLH on the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (red) and the $t\bar{t}$+jets background samples (blue).



(a) E_H^{KL} (reco kB6) (b) E_H^{KL} (reco kB7) (c) E_H^{KL} (reco kB8)

(d) E_H^{KL} (matched-reco kB6) (e) E_H^{KL} (matched-reco kB7) (f) E_H^{KL} (matched-reco kB8)

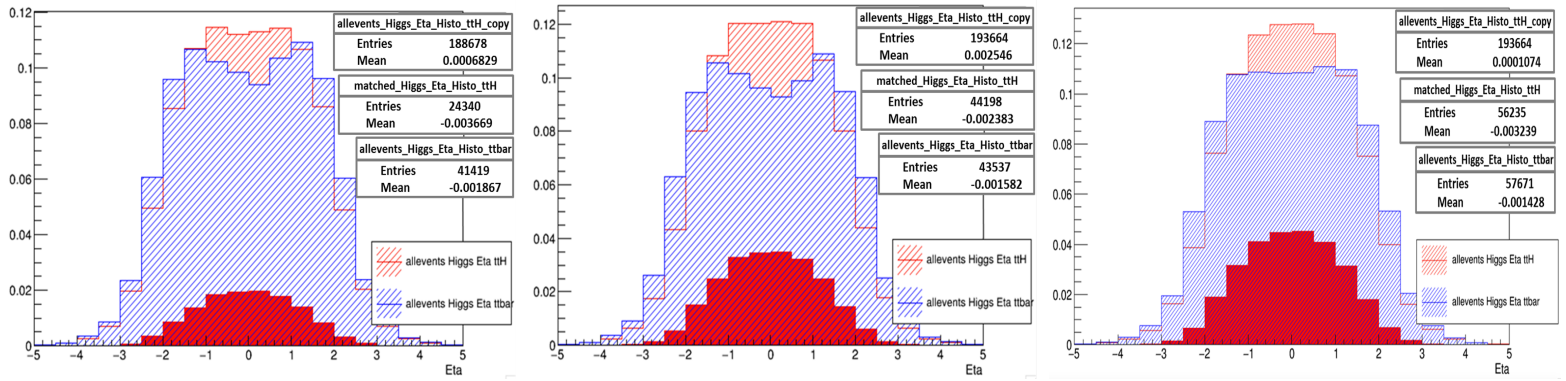
Figure 8.22: Upper row: the impact of using the kB6, kB7 and kB8 modes of KLFitterLH on the normalized reconstructed Higgs boson E [GeV] distributions, for $t\bar{t}$+jets (blue) and $t\bar{t}H$ (red) signal samples. Lower row: illustration of the effect on the shape of the reconstructed Higgs boson E distributions, when applying the kB6, kB7 and kB8 modes of KLFitterLH on the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (red) and the $t\bar{t}$+jets background samples (blue).



(a) η_H^{KL} (reco kB6) (b) η_H^{KL} (reco kB7) (c) η_H^{KL} (reco kB8)

(d) η_H^{KL} (matched-reco kB6) (e) η_H^{KL} (matched-reco kB7) (f) η_H^{KL} (matched-reco kB8)

Figure 8.23: Upper row: the impact of using the kB6, kB7 and kB8 modes of KLFitterLH on the normalized reconstructed Higgs boson η distributions, for $t\bar{t}$+jets (blue) and $t\bar{t}H$ (red) signal samples. Lower row: illustration of the effect on the shape of the reconstructed Higgs boson Eta distributions, when applying the kB6, kB7 and kB8 modes of KLFitterLH on the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (red) and the $t\bar{t}$+jets background samples (blue).



(a) η_H^{KL} (matched-reco kB6) (b) η_H^{KL} (matched-reco kB7) (c) η_H^{KL} (matched-reco kB8)

Figure 8.24: Illustration of the effect on the shape of the reconstructed Higgs boson η distributions, when applying the kB6, kB7 and kB8 modes of KLFitterLH on the $t\bar{t}$+jets background (blue), $t\bar{t}H$ signal (red), and the truth-matched ($\Delta R_{\text{truth-reco}} \leq 0.5$) portion of the $t\bar{t}H$ signal (solid red).

Investigation of efficiency with increased jet multiplicities (7 and 8 jets modes)

The overall increase in particle reconstruction efficiency when more jets are included in KLFitterLH is now investigated. As mentioned in the previous Section, the truth-matching efficiency for the Higgs, tlep and thad and both tops together in signal is the highest for kB8 then kB7 then finally kB6 mode. The questions regarding these results are:

- Is it possible that the Higgs boson was not reconstructed properly because the jets which belonged to the Higgs was not handed over to KLFitterLH?
- Are the b-tags that belong to Higgs or thad and tlep sometimes at low p_T where they would be in the 7th or 8th psition?
- Are the light jets belonging to the thad excluded, even though they have high p_T (because of the kB mode)?

To investigate these questions, the jets are first truth-matched to the Higgs, thad and tlep truth quarks. Plots are then produced showing the position ordered which these jets occupy, both for unmatched and successfully truth-matched events. Perhaps with only 6 jets KLFitterLH does not have the correct jets available to reconstruct the Higgs or tops? First, the ΔR between each jet in the event and the truth Higgs is calculated. Then, the two jets having the lowest ΔR to the truth Higgs (which may be the two b-tags from the Higgs) are found and binned according to the p_T ordering from zero (highest p_T) to nine (lowest p_T), see Figures 8.25a, 8.25a and 8.25c. This process is repeated for events where the reconstructed Higgs boson matches truth, see Figures 8.25d, 8.25e and 8.25f. The study is repeated for thad and tlep, but by looking for the lowest $\Delta R_{(\mathrm{truth}_{top}-\mathrm{reco}_{jet})}$ of three jets for thad, and the lowest $\Delta R_{(\mathrm{truth}_{top}-\mathrm{reco}_{jet})}$ of one jet for tlep. For the Higgs study, in all KLFitterLH selection modes (6-7-8 jets) when we consider all events, see Figures (8.25a), (8.25b) and 8.25c, then the jets that are matched to the reconstructed Higgs often the second and third leading jets (jet 1 and jet 2). In the case where the truth Higgs is matched to the reconstructed Higgs from KLFitterLH displayed in Figures 8.25d, 8.25e and 8.25f, the jets that belong to the Higgs are often the leading or subleading jets (jet0 and jet1). When kB7 or kB8 mode is chosen, the increase of the efficiency can be explained by the following. On the b-tag side, the KLFitterLH is able to use more jets in the hope that two b's from the Higgs or top decays will be included (more often). For light jets, kB choice means keeping all b-jets that it

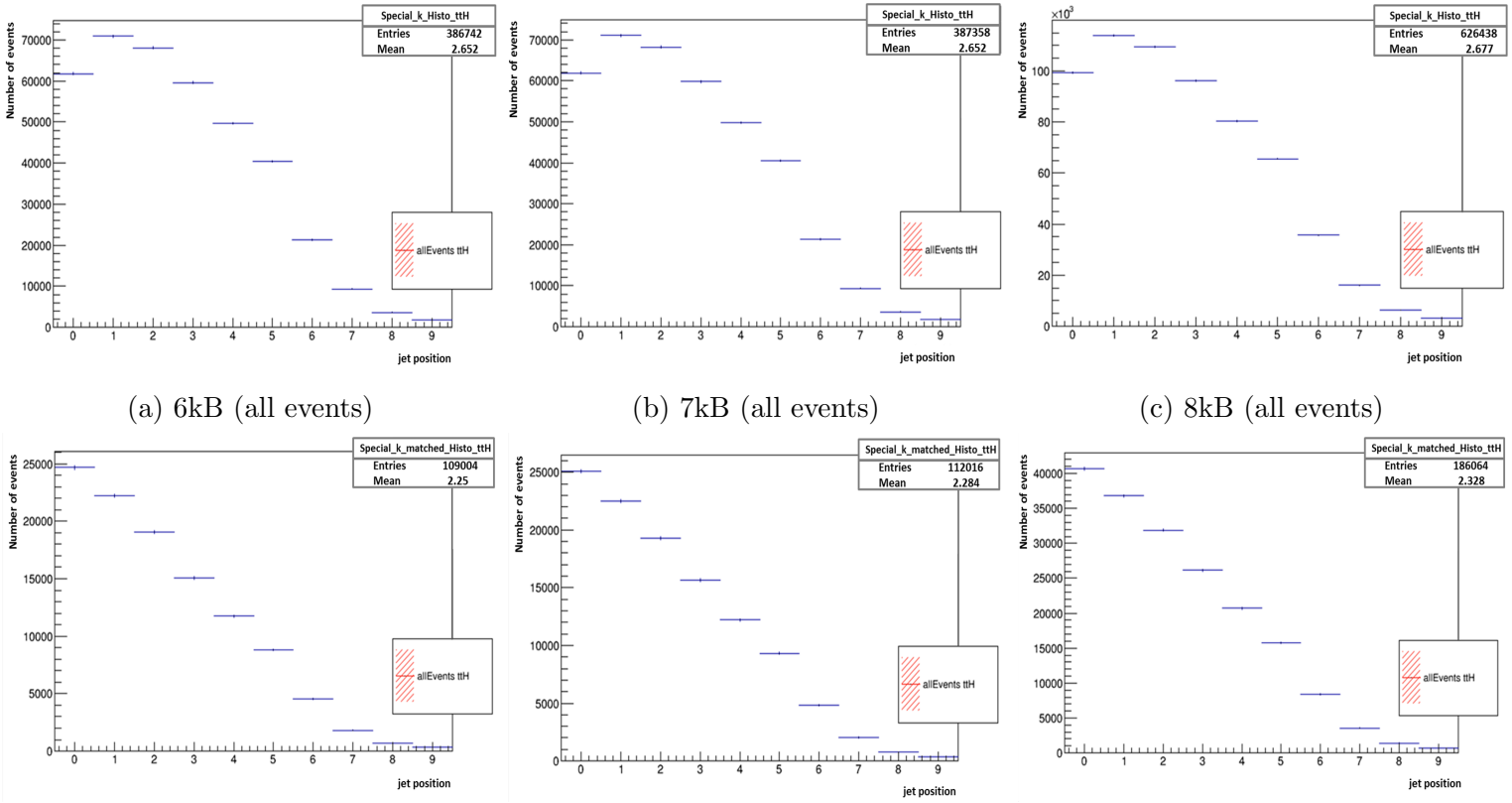
can find, even to the exclusion of some important (high p_T) light jets. Using 7 or 8 jet selection modes means that b-tags are allowed in, but there is more space for KLFitterLH to accept light jets (some with high p_T) as well.

As for thad, Figures 8.26a to 8.26f, in general show the shapes of all the plots are very similar. There is a plateau across positions 0-2 in signal events in general and matched events, showing that the jets used in thad are typically in p_T-ordered positions 0-2. The only noticeable difference looks to be in the kB6 mode for events with a truth-matched thad (Figure 8.26d), where the leading jet (position 0) is used more often, to the detriment of the other jets in positions in 1,2,3,4,$\cdots$. For the tlep study, Figures from 8.27a to 8.27f show that the shapes are very similar for all six plots. The tlep jet is typically located in the second p_T-ordered slot, but there is a tendency for the tlep jet to be in the first few (0-2) slots (i.e it has high p_T). When only *matched* events are considered (bottom row) there is a slightly greater tendency for the tlep jet to be in the second position compared to the first.

In general, the matched events have better assignment by definition, because they are matched. Typically, jets from tlep, thad or Higgs will have high p_T compared to ISR/FSR/fake jets. Therefore, when tlep, thad and Higgs are matched, the end result is that the jets which are assigned have a higher p_T.

8.3 Summary

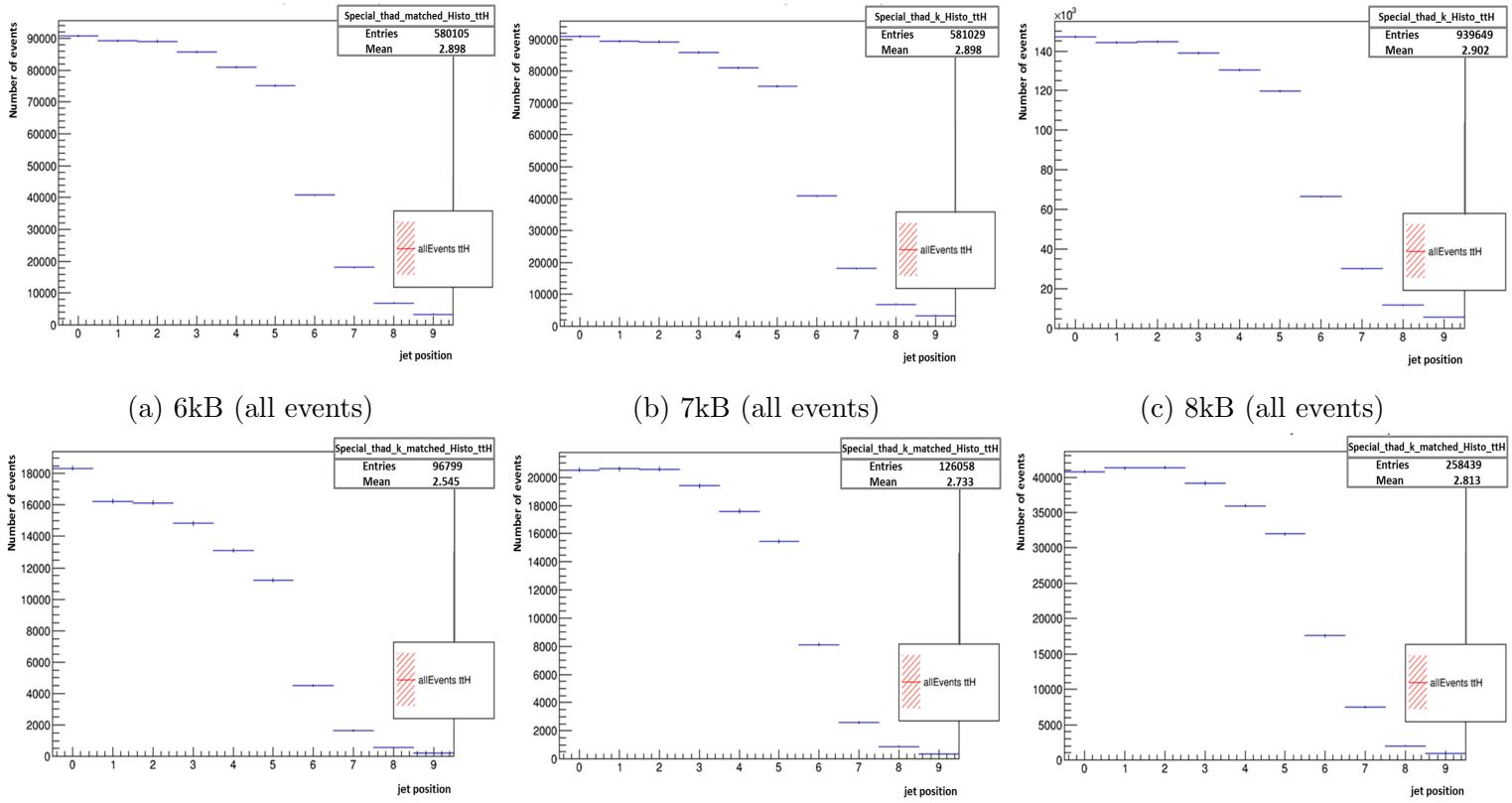
As is detailed, the five variables which are provided by the KLFitterLH do have some discrimination power between the signal and background samples after three stages of reduction. Furthermore, the discriminating power of these variables is tested in three KLFitterLH selection modes kB6, kB7 and kB8. It is found that the Higgs, tlep and thad and both tops together in signal have the highest efficiency under the kB8 mode. The investigation of efficiency with increased jet multiplicities, kB7 and kB8, is discussed to show that the efficiency increased in these two modes by providing an extra one(two) slots for the b-tags in the hope that two bs from the Higgs or top decays will be included. The increase of the efficiency can also be explained from a light-jets perspective. Some important (high p_T) light jets are included which were excluded in kB6 because of the use of kB, which means keeping all b-jets that it can find. It is possible to exploit the discriminating power of the five KLFitterLH variables by using a more advanced technique which is applied in the next chapter by combining the KLFitterLH variables with the standard variables in the MultiVariate Analysis (BDTs).



(a) 6kB (all events) (b) 7kB (all events) (c) 8kB (all events)

(d) 6kB (matched events) (e) 7kB (matched events) (f) 8kB (matched events)

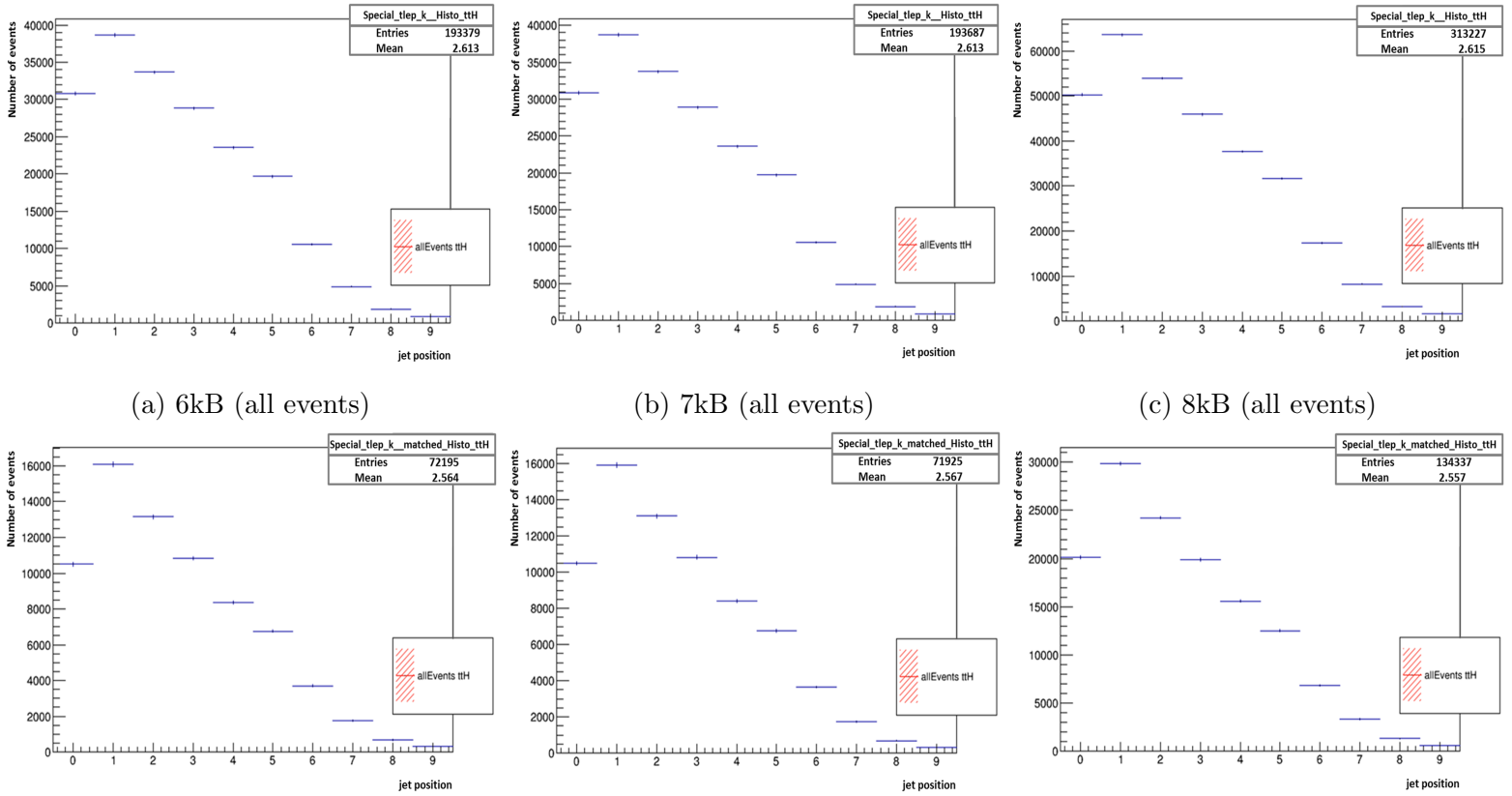
Figure 8.25: Investigation of efficiency in the truth-matched Higgs boson (bottom) and unmatched (top) with increased jet multiplicities (7 and 8 jets modes). No differences can be seen between the modes.



(a) 6kB (all events) (b) 7kB (all events) (c) 8kB (all events)

(d) 6kB (matched events) (e) 7kB (matched events) (f) 8kB (matched events)

Figure 8.26: Investigation of efficiency in the truth-matched hadronic top quark (bottom) and unmatched (top) with increased jet multiplicities (7 and 8 jets modes). The addition of extra slots to allow more jets into the KLFitter permutations significantly affects the composition of the hadronic top, most noticeably in the 6kB to 7kB transition. The light jets in the positions 2 and 3 are often discarded in the 6kB, but are much more likely to be included when using the 7kB and 8kB modes.



(a) 6kB (all events) (b) 7kB (all events) (c) 8kB (all events)

(d) 6kB (matched events) (e) 7kB (matched events) (f) 8kB (matched events)

Figure 8.27: Investigation of efficiency in the truth-matching leptonic top (bottom) and unmatched (top) quark with increased jet multiplicities (7 and 8 jets modes). No differences can be seen between the modes.

Chapter 9

Signal strength measurement

In this chapter, the measurement of the observed signal strength of the $t\bar{t}H$ production cross section is presented. The statistical limit that is associated with the calculation of the signal strength is also provided. These results were obtained from the analysis of the 2016 data samples corresponding to 32.9 fb^{-1} of data. A comparison is made between *multivariate analysis* (MVA) using non-KLFitter variables and the combined KLFitter and non-KLFitter variables is giving. Statistical and systematic uncertainties are introduced in Sections 9.2 and 9.3, respectively, in order to derive limits on the signal strength parameter in each case.

9.1 KLFitter and non-KLFitter results

9.1.1 BDT training of the signal and $t\bar{t}$+jets background

As mentioned in Section 4.2, *Boosted Decision Trees* (BDTs) are used to discriminate potential signal events from the background in the full reconstruction of the $t\bar{t}H(H \to b\bar{b})$ semileptonic system. They are particularly useful in cases where no single variable exhibits a clear separation power between signal and background. BDTs allow the combination of information from several input variables into one output discriminant that exploits the correlations among the variables and can reproduce a non-trivial selection in the variables phase space.

The study commences by training the events in the inclusive (6j,4b) region, using two Monte-Carlo $t\bar{t}H$ and two $t\bar{t}$+jets samples, at $\sqrt{s}$=13 TeV. This study uses 10

of the same variables (see Table 9.1)from a previous study [7], which had a total of 12 variables. This analysis also increases the center-of-mass to 13 TeV, compared to 8 TeV in [7]. Whilst this part of the study is not intended to replicate [7], it provides a baseline for further investigation. As the study does not yet include KLFitter variables, it is named the "non-KLFitter" analysis. Many variations of the BDT training parameters were tried for the BDT but none led to noticeable improvements except the ones that are shown in Table 9.2. The training results are shown in terms of BDT response, separation and significance in the Table 9.3 and Figure 9.1.

Variable	Description
Centrality	Scalar sum of the p_T divided by sum of the E for all jets and the lepton
H1	FoxWolfram moments describe the geometrical correlation among objects in the event, in terms of spherical harmonics [136]. The Second Fox-Wolfram moment variable which is computed for all jets and the lepton in the event, is used in this analysis.
p_T^{jet5}	p_T of the fifth leading jet
ΔR_{bb}^{avg}	Average ΔR for all b-tagged jet pairs
Aplan_{b-jets}	Aplanarity is defined as 1.5 times the second eigenvalue of the momentum tensor (λ_2) [137]. It is used to look at different event topologies by distinguishing the spherical from planar and linear events. It is built only with the b-tagged jets.
$m_{bj}^{\text{max}p_T}$	Mass of the combination of a b-tagged jet and any jet with the largest vector sum p_T
$m_{bb}^{\text{min}\Delta R}$	Mass of the combination of the two b-tagged jets with the smallest ΔR
$\Delta R_{bb}^{\text{max}p_T}$	ΔR between the two b-tagged jets with the largest vector sum p_T
$\Delta R_{lep-bb}^{\text{min}\Delta R}$	Minimum ΔR between the lepton and any pair of b-tagged jets
$m_{uu}^{\text{min}\Delta R}$	Mass of the combination of the two untagged jets with the smallest ΔR

Table 9.1: List of the input nonKFitter variables (10 variables) that are used in the training of the events in the inclusive (6j,4b) single lepton region.

Subsequently, the 5 outputs from KLFitterLH, see Table 8.3, are added to the

[1]D1,NeymanPearson MEM discriminant, and SSLL, Logarithm of the summed signal likelihoods are not included in study.

TMVA setting	$\geq 6j, \geq 4b$
NTrees	500
MaxDepth	2
nCuts	20
MinNodeSize	2%

Table 9.2: BDT parameters.

Variables	ROC-integral	Separation	Significance
10 ranked vars	0.719	0.147	0.583

Table 9.3: ROC curve integral, separation and significance when using 10 non-KLFitter variables.

previous 10 variables by including each one at a time, to see the effect, see Table 9.4. Additionally, the kB6 jet selection mode of KLFitter is used for the first time. It is clear that the significance and the separation power increase, and the most powerful variable is the Higgs mass. The correlations among the variables (the KLFitter and non-KLFitter variables) for signal and background are shown in Figure 9.3 and the BDT response is displayed in Figure 9.2. The ranked separation power and the overall importance are displayed in Table 9.5. These are the factors taken into account in order to choose the final set of discriminating variables. The Higgs boson energy variable is highly correlated with the two KLFitterLH variables, Higgs mass and Higgs p_T, thus, it is removed and the significance and the separation sensitivity was not affected as displayed in Table 9.6.

The study is then extended by training using the combined 10 non-KLFitter and the 4 KLFitter variables for the same samples in two new modes: the kB7 and kB8 KFitter jet selection modes. No appreciable improvement in the significance and separation sensitivity is observed, the largest effect being of the order of $\simeq 1\%$, see Table 9.7. This difference is neglected in the TMVA study, and the rest of the analysis will continue with the kB6 KLFitter jet selection mode.

BDT training of the signal and all backgrounds.

Up to this point, the analysis has used one signal region ($\geq 6j, \geq 4b$), and trained two BDTs: non-KLfitter variables, and non-KLFitter+KLfitter variables. This Section

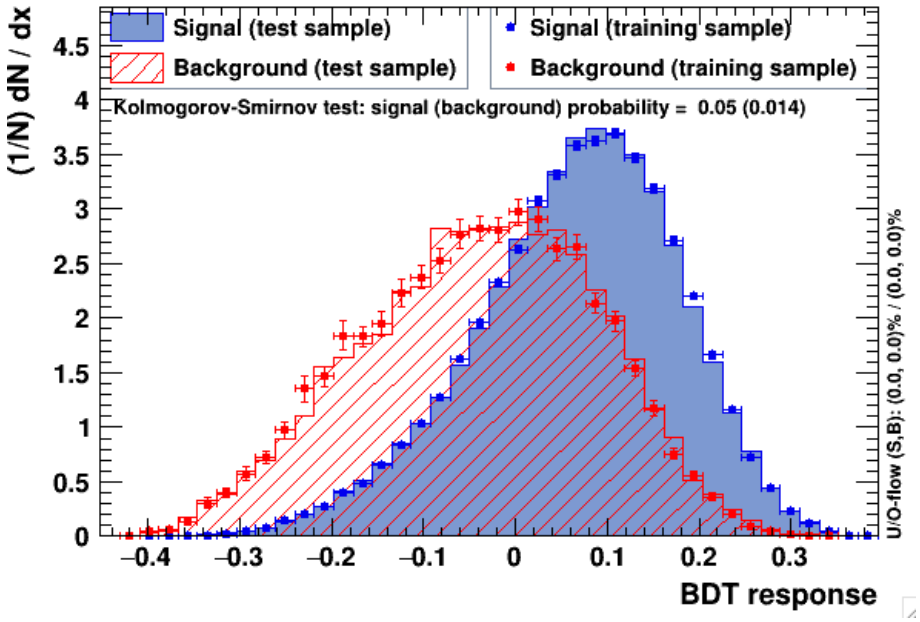


Figure 9.1: Classification-BDT distribution for training and testing samples when the 10 non-KLFitter variables are used.

expands the study to three regions ($\geq 6j, \geq 4b$), ($\geq 6j, 3b$) and ($5j, \geq 4b$), and uses all available background samples, plus the signal.The samples are those described in Section 6.2. The centre-of-mass energy also increases from 8 TeV to 13 TeV, and consequently, the set of non-KLFitter variables are slightly different, depending upon the region considered, see Table 9.8.

There are four BDTs used in this Section; three trained in the regions ($\geq 6j, \geq 4b$), ($\geq 6j, 3b$) and ($5j, \geq 4b$) using only non-KLFitter variables, and one BDT trained in the ($\geq 6j, \geq 4b$) region, using the combined non-KLFitter+KLFitter variables. There are two BDTs trained in the ($\geq 6j, \geq 4b$) region. The first uses a non-KLFitter kinematic variables based on ICHEP results [61], and the other BDT uses these 9 non-KLFitter variables, plus 4 KLFitter variables, as shown in Table 9.8. The last two BDTs use only non-KLFitter variables in the two regions ($\geq 6j, 3b$) and ($5j, \geq 4b$) as per the ICHEP [61]. BDT parameters that are used in all four BDTs are shown in Table 9.9. The complete set of distributions (signal and background shapes) of all the 13 variables in the region ($\geq 6j, \geq 4b$) are shown in Figures 9.4 and 9.6, in order to evaluate the discrimination power. Table 9.10 reports the ranked separation power and overall importance for the second iteration of the BDT with the 13 variables. The final correlations among the 13 variables in signal and background are shown in Figure 9.7.

KLFitterLH is only used in the ($\geq 6j, \geq 4b$) region. Whilst a BDT is used in the

Variables	ROC-integral	Separation	Significance
10 ranked vars	0.719	0.147	0.583
m_H^{KL}	0.724	0.153	0.597
E_H^{KL}	0.722	0.150	0.590
$p_{T,H}^{KL}$	0.720	0.148	0.587
$\Delta R_{b1b1,H}^{KL}$	0.719	0.147	0.584
η_H^{KL}	0.718	0.145	0.581
adding all the KLFitter vars	0.725	0.154	0.602

Table 9.4: ROC curve integral, separation and significance when using 10 non-KLFitter variables and including only one KLFitterLH variable at a time, with the previous 10 variables to see the effect of each variable.

Ranking	Variable	Separation power
1	ΔR_{bb}^{avg}	0.0636
2	p_T^{jet5}	0.0403
3	$\Delta R_{bb}^{\max p_T}$	0.0371
4	Centrality	0.0304
5	$m_{bb}^{\min \Delta R}$	0.0267
6	m_H^{KL}	0.0259
7	Aplan_{b-jets}	0.0255
8	$\Delta R_{b1b2,H}^{KL}$	0.0215
9	$\Delta R_{lep-bb}^{\min \Delta R}$	0.0163
10	$m_{bj}^{\max p_T}$	0.0094
11	η_H^{KL}	0.0076
12	E_H^{KL}	0.0066
13	$H1$	0.0058
14	$p_{T,H}^{KL}$	0.0058
15	$m_{uu}^{\min \Delta R}$	0.0024

(a)

Ranking	Variable	Importance
1	ΔR_{bb}^{avg}	0.0976
2	Centrality	0.0921
3	H1	0.0823
4	$\Delta R_{bb}^{\max p_T}$	0.0779
5	p_T^{jet5}	0.0765
6	$\Delta R_{lep-bb}^{\min \Delta R}$	0.0765
7	η_H^{KL}	0.0716
8	$m_{bb}^{\min \Delta R}$	0.0681
9	Aplan_{b-jets}	0.0660
10	m_H^{KL}	0.0618
11	$p_{T,H}^{KL}$	0.0609
12	$m_{bj}^{\max p_T}$	0.0486
13	E_H^{KL}	0.0486
14	$\Delta R_{b1b1,H}^{KL}$	0.0374
15	$m_{uu}^{\min \Delta R}$	0.0343

(b)

Table 9.5: Separation power (left) and importance (right) of the 15 non-KLFitter and KLFitter input variables. Variables are defined in Table 9.1.

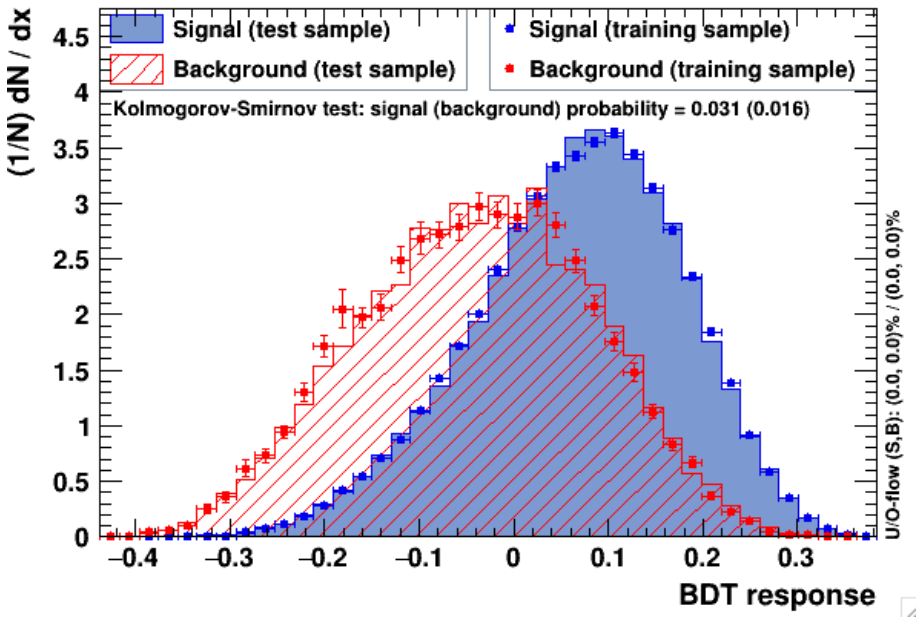


Figure 9.2: Classification-BDT distribution for training and testing samples when the 15 non-KLFitter and KLFitter variables are used.

Variables	Roc-integ	Separation	Significance
All KLFitter vars	0.725	0.154	0.602
Remove E_H^{KL}	0.726	0.155	0.602

Table 9.6: ROC curve integral, separation and significance when using 10 non-KLFitter plus 5 KLFitter variables in the first row, and the effect of removing the Higgs energy variable in the second row.

signal-rich regions ($\geq 6j, \geq 4b$), ($\geq 6j, 3b$) and ($5j, \geq 4b$), the analysis does not use BDTs in the signal-depleted regions: (4 jets, 2 b-tags), (4 jets, 3 b-tags), (4 jets, 4 b-tags), (5 jets, 2 b-tags), (5 jets, 3 b-tags) and (≥6 jets, 2 b-tags). Consequently, in the fit which will be used, the variable of choice for the signal-depleted regions cannot be the BDT weight; the variable used instead is H_T^{had} (the scalar sum of jet p_T, see Table 9.1). This variable is chosen due to its insensitivity to the systematic uncertainties such as jet energy scale (JES) or b-tagging, which have a clear p_T dependence, and to the background modeling. The signal-depleted regions have high data statistics and the fit of H_T^{had} allows control of the impact of systematic uncertainties primarily affecting $t\bar{t}$+jets events. The comparison of data and MC predictions for the H_T^{had} and BDT distributions in each of the analysis channels considered are shown in Figures 9.8 9.9 and 9.10.

In summary, the MVA analyses (BDTs from TMVA) use the variables from Table

KLF mode	kB6	kB7	kB8
Significance	0.602	0.605	0.610
Separation	0.155	0.156	0.158
ROC-integral	0.726	0.727	0.729

Table 9.7: ROC integral, separation and significance when using the KLFitter variables in kB6, kB7 and kB8 jet selection modes.

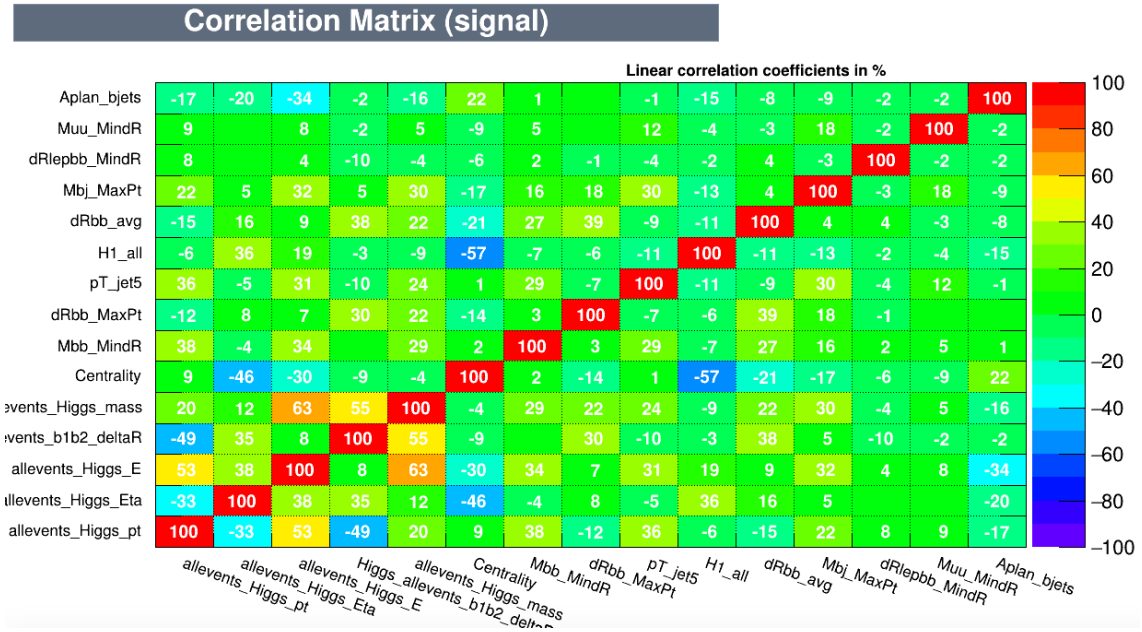
8.3 as inputs to discriminate signal from background giving each event a BDT weight to indicate how signal-like it is. In the end, the BDT training using KLFitter+non-KLFitter variables is selected and together with the other two BDTs (non-KLFitter regions ($\geq 6j, 3b$) and ($5j, \geq 4b$)), plus H_T^{had} from the remaining six signal-depleted regions are all used to set the upper limit $\mu = \frac{\sigma_{\text{observed}}}{\sigma_{\text{expected}}}$ (see Section 4.3). This is accomplished by fitting the shape of the MVA output (BDT weight) plus the H_T^{had} distributions using the TTFitter package, as used by the ATLAS $t\bar{t}H$ group to perform the fit to data.

Variable	Description	Region			
		($\geq 6j, \geq 4b$) klf+non-klf	($\geq 6j, \geq 4b$) non-klf	($\geq 6j, 3b$)	($5j, \geq 4b$)
$\Delta Rb1b1_H^{KL}$		✓	-	-	-
m_H^{KL}		✓	-	-	-
$p_{T,H}^{KL}$		✓	-	-	-
E_H^{KL}		✓	-	-	-
Centrality		✓	✓	✓	✓
$\Delta\eta_{jj}^{\max\ \Delta\eta}$		✓	✓	✓	✓
H1		✓	✓	✓	✓
p_T^{jet5}		✓	✓	✓	✓
ΔR_{bb}^{avg}		✓	✓	✓	✓
Aplan	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor built with all jets	✓	✓	✓	✓
N_{30}^{Higgs}	Number of b-jet pairs with invariant mass within 30 GeV of the Higgs boson mass	✓	✓	-	✓
$m_{bb}^{\min\Delta R}$		✓	✓	✓	-
$m_{bj}^{\max p_T}$		-	-	✓	-
$\Delta R_{bb}^{\max p_T}$		✓	✓	-	-
$\Delta R_{lep-bb}^{\min\Delta R}$		-	-	-	✓
N_{40}^{jet}	Number of jets with $p_T \geq$40 GeV	-	-	✓	-
H_T^{had}	Scalar sum of jet	-	-	✓	✓
$m_{jj}^{\min\Delta R}$	Mass of the combination of any two jets with the smallest ΔR	-	-	-	✓

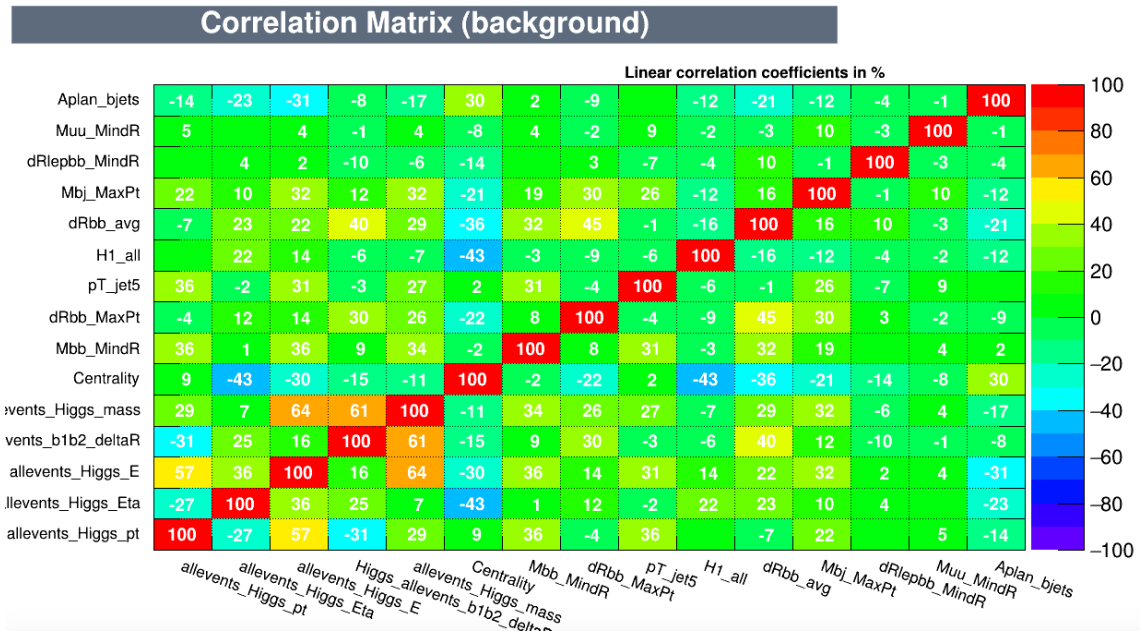
Table 9.8: List of the input variables for the BDTs classifier in the single lepton region. The undefined variables are previously defined in Tables 9.1 and 8.3.

TMVA setting	Value
NTrees	400
MaxDepth	3
nCuts	40
MinNodeSize	4%

Table 9.9: Classification-BDT parameters used in all four BDTs for the second iteration (13 TeV, 13 variables).



(a)



(b)

Figure 9.3: Correlations amongst input variables in the initial set for the classification-BDT training for signal (a) and background (b).

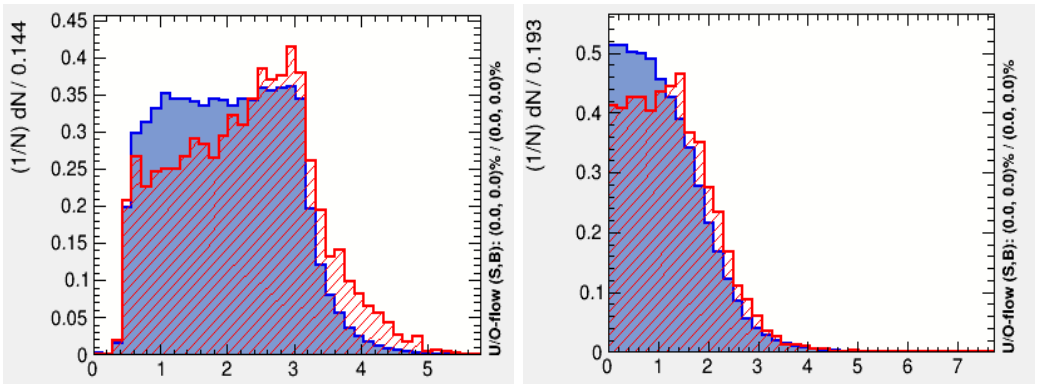
Ranking	Variable	Separation power
1	ΔR_{bb}^{avg}	0.0734
2	$\Delta\eta_{jj}^{\max\Delta\eta}$	0.0442
3	Centrality	0.0369
4	$\Delta R_{bb}^{\max p_T}$	0.0368
5	N_{30}^{Higgs}	0.0334
6	p_T^{jet5}	0.0276
7	$\Delta R_{b1b2,H}^{KL}$	0.0270
8	Aplan	0.0267
9	$m_{bb}^{\min\Delta R}$	0.0226
10	m_H^{KL}	0.0193
11	$p_{T,H}^{KL}$	0.0106
12	η_H^{KL}	0.0105
13	$H1$	0.00518

(a)

Ranking	Variable	Importance
1	ΔR_{bb}^{avg}	0.143
2	Centrality	0.0878
3	$m_{bb}^{\min\Delta R}$	0.0861
4	p_T^{jet5}	0.0836
5	H1	0.0832
6	N_{30}^{Higgs}	0.0764
7	$\Delta\eta_{jj}^{\max\Delta\eta}$	0.0737
8	$\Delta R_{bb}^{\max p_T}$	0.0689
9	Aplan	0.0671
10	m_H^{KL}	0.0606
11	$p_{T,H}^{KL}$	0.0570
12	$\Delta R_{b1b2,H}^{KL}$	0.0562
13	η_H^{KL}	0.0483

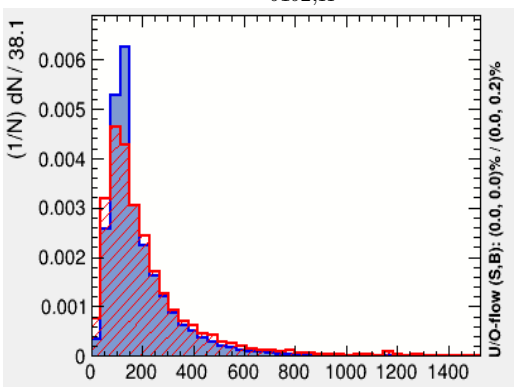
(b)

Table 9.10: Separation power (left) and importance (right) of the 13 non-KLFitter and KLFitter input variables (Variables defined in Table 9.1).

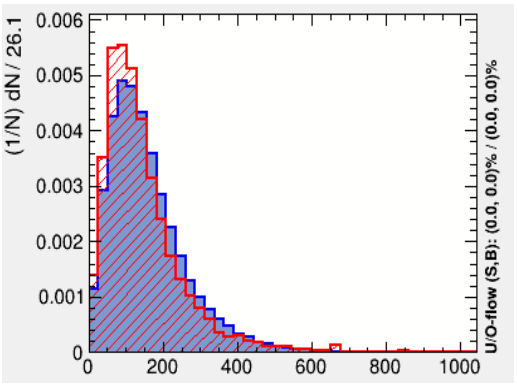


(a) $\Delta R^{KL}_{b1b2,H}$

(b) η^{KL}_{H}

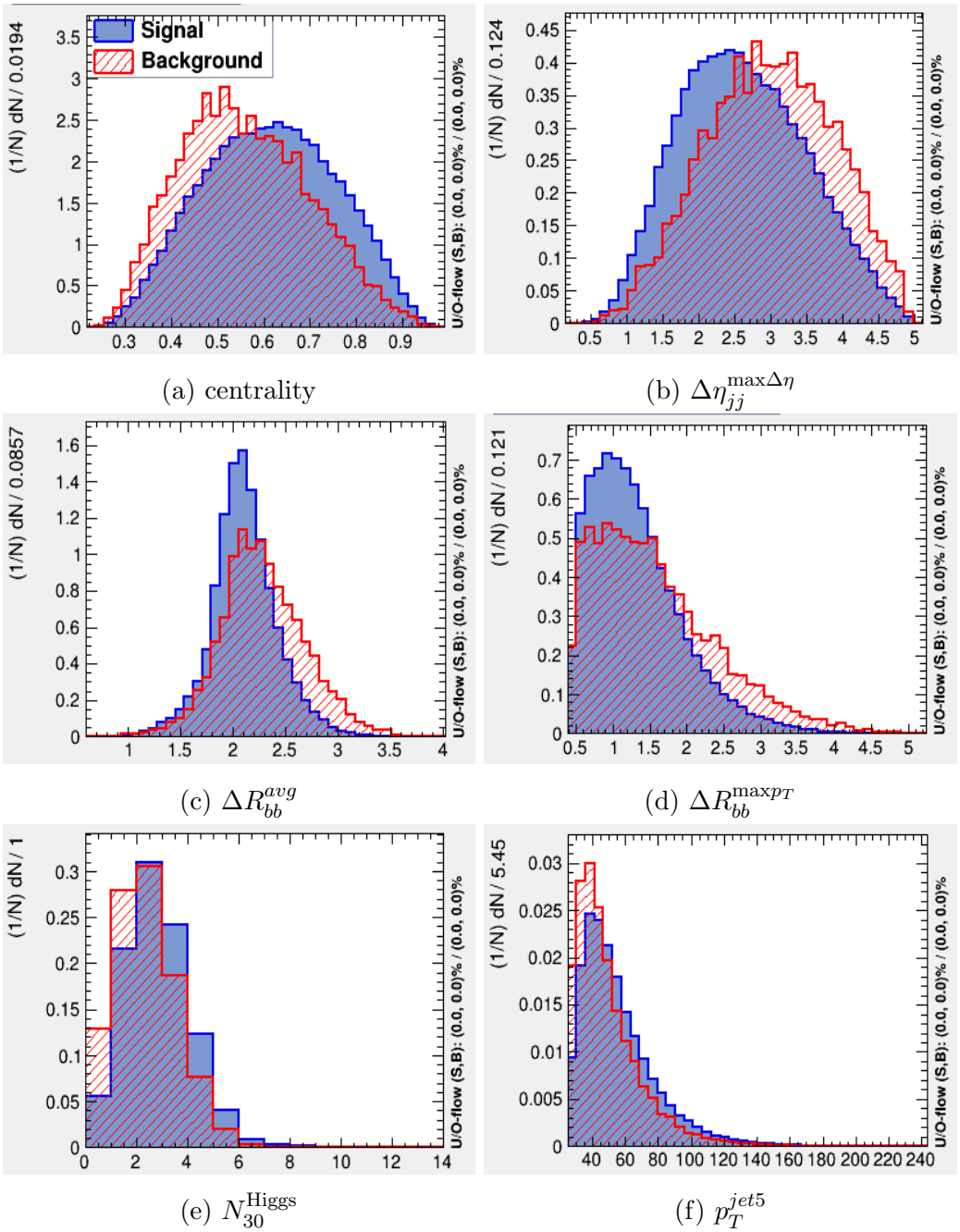


(c) m^{KL}_{H}



(d) $p^{KL}_{T,H}$

Figure 9.4: Distribution of signal (blue) and background (red) for the four input KLFitter variables to the BDT, in ($\geq 6j, \geq 4b$) inclusive signal region.

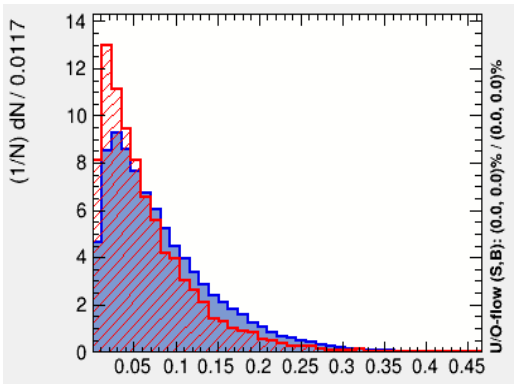


(a) centrality (b) $\Delta\eta_{jj}^{\max\Delta\eta}$

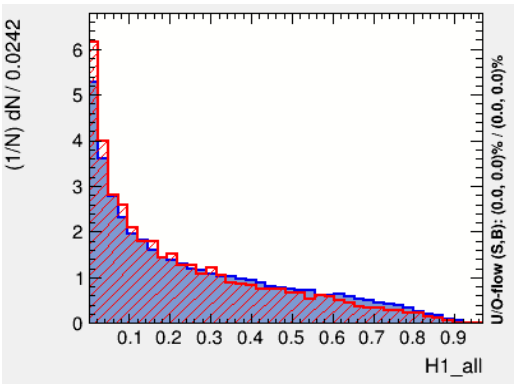
(c) ΔR_{bb}^{avg} (d) $\Delta R_{bb}^{\max p_T}$

(e) N_{30}^{Higgs} (f) p_T^{jet5}

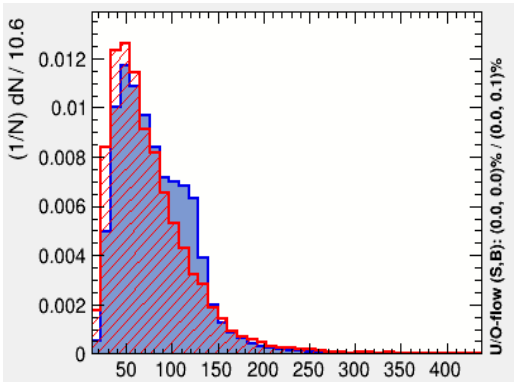
Figure 9.5: Distribution of signal (blue) and background (red) for the nine input non-KLFitter variables to the BDT, in the ($\geq 6j, \geq 4b$) inclusive signal region.



(a) Aplan

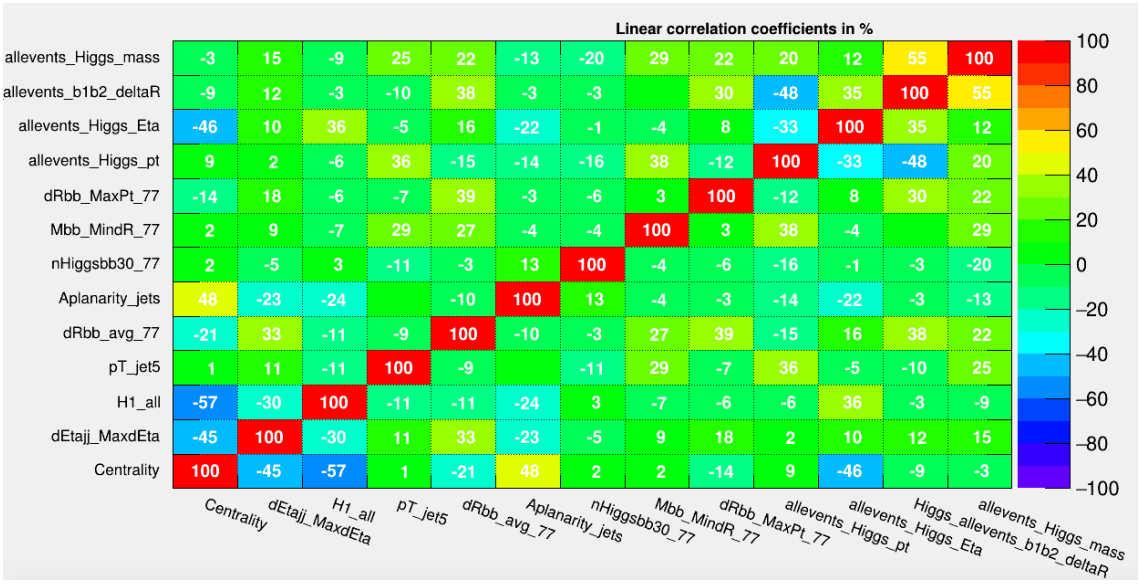


(b) H1

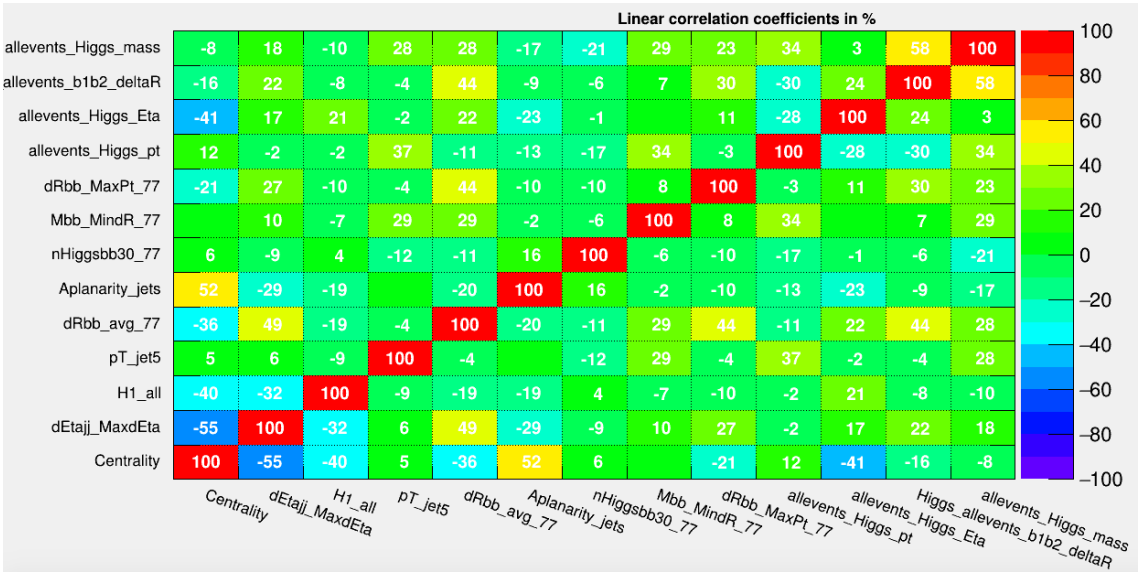


(c) $m_{bb}^{\min\Delta R}$

Figure 9.6: Continued, distribution of signal (blue) and background (red) for the nine input non-KLFitter variables to the BDT, in the ($\geq 6j, \geq 4b$) inclusive signal region.

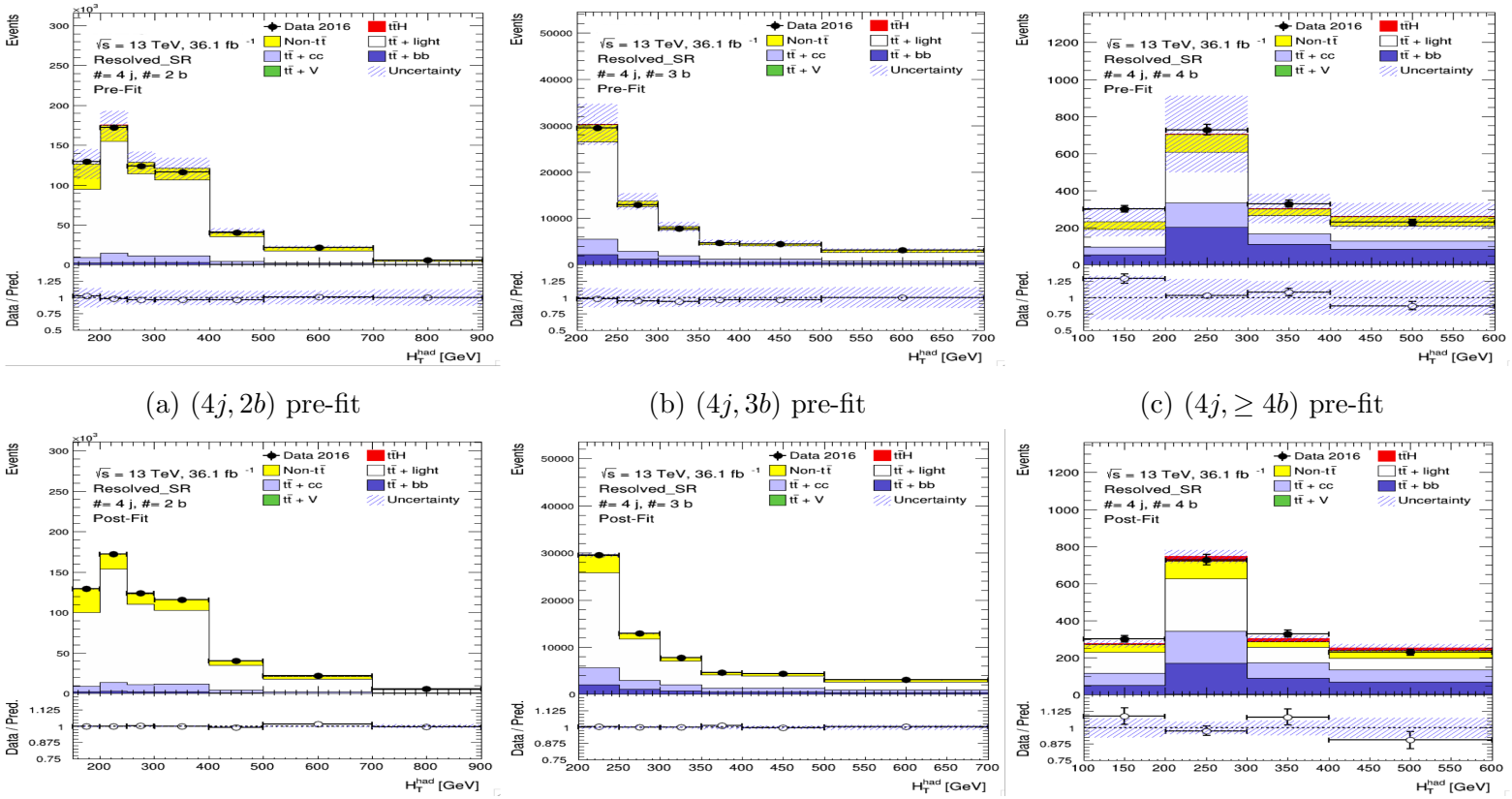


(a)



(b)

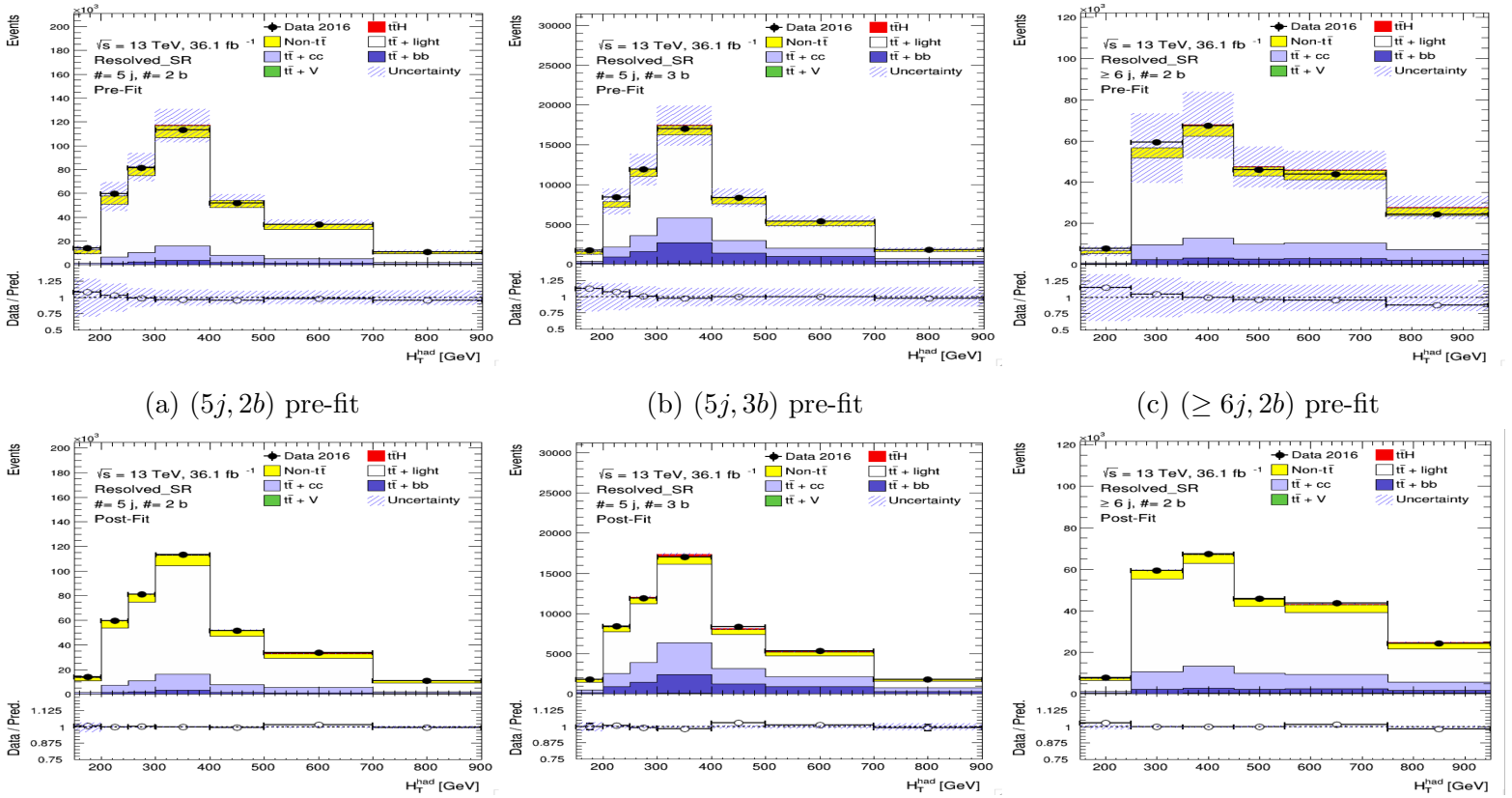
Figure 9.7: Correlations amongst input variables in the second iteration (13 TeV, 13 variables) for the classification-BDT training for (a) signal and (b) background.



(a) $(4j, 2b)$ pre-fit (b) $(4j, 3b)$ pre-fit (c) $(4j, \geq 4b)$ pre-fit

(d) $(4j, 2b)$ post-fit (e) $(4j, 3b)$ post-fit (f) $(4j, \geq 4b)$ post-fit

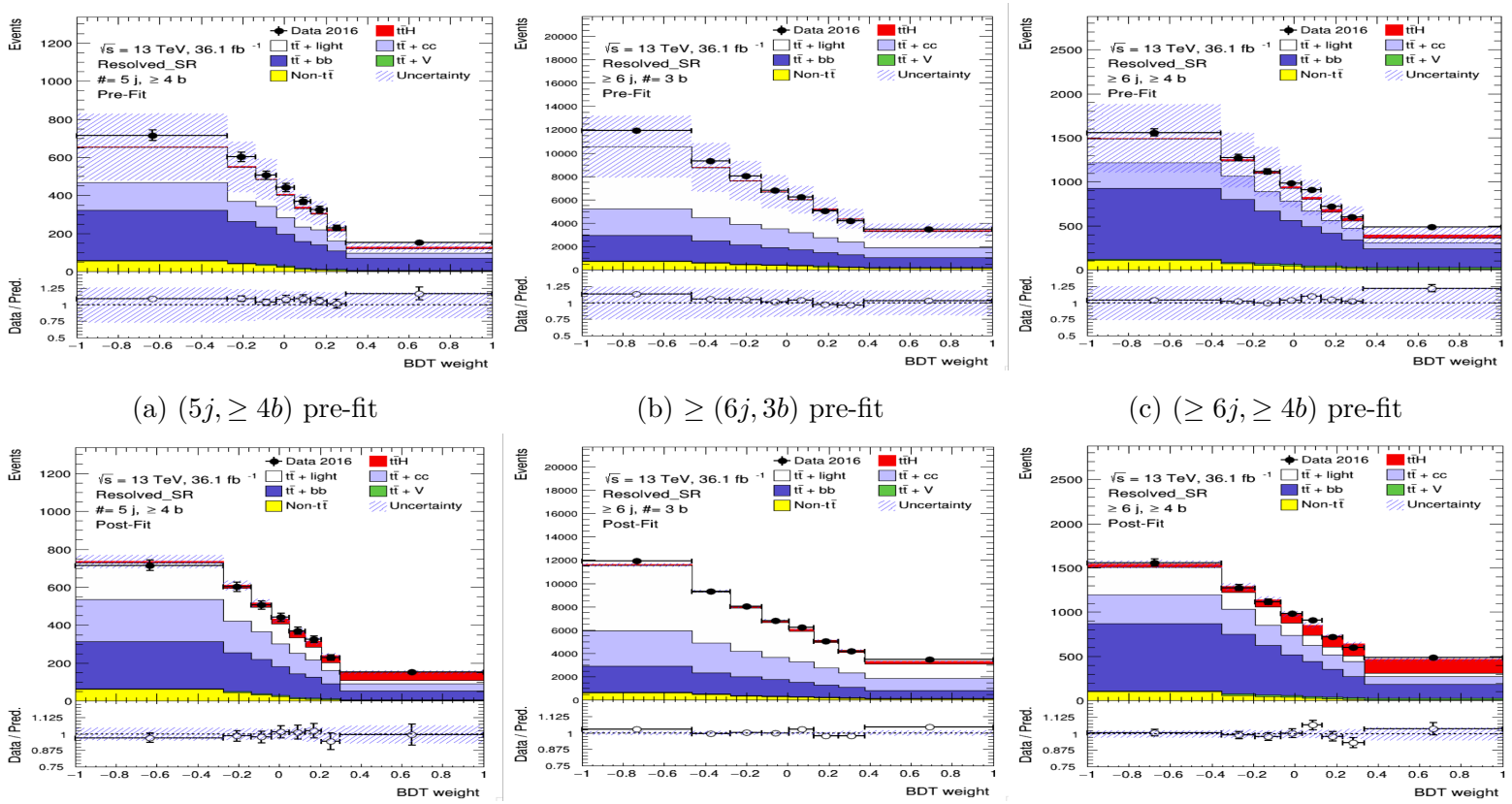
Figure 9.8: Comparison between the data and KLFitter prediction for the H_T^{had} distributions before and after performing the fit to data.



(a) $(5j, 2b)$ pre-fit (b) $(5j, 3b)$ pre-fit (c) $(\geq 6j, 2b)$ pre-fit

(d) $(5j, 2b)$ post-fit (e) $(5j, 3b)$ post-fit (f) $(\geq 6j, 2b)$ post-fit

Figure 9.9: Continued, comparison between the data and KLFitter prediction for the H_T^{had} distributions before and after performing the fit to data.



(a) $(5j, \geq 4b)$ pre-fit (b) $\geq (6j, 3b)$ pre-fit (c) $(\geq 6j, \geq 4b)$ pre-fit

(d) $(5j, \geq 4b)$ post-fit (e) $(\geq 6j, 3b)$ post-fit (f) $(\geq 6j, \geq 4b)$ post-fit

Figure 9.10: Comparison between the data and KLFitter prediction for the classification-BDT output distributions before and after performing the fit to data.

9.2 Statistical uncertainties

The statistical uncertainty originates from the finite number of both selected data events and simulated MC events. The MC events can be produced in order to provide predictions for real data. For this reason, the measurement precision is affected mainly by the limited data statistics. The statistical uncertainty in each bin is calculated by taking the the square root of the observed number of events. It is then summed in quadrature with the total systematic uncertainty, since the two can to be considered uncorrelated. The fit of the distribution in each bin is done using H_T^{had} for the control regions and the BDT output for the signal regions. Finally, the fit procedure described in Section 4.3 determines the final value and the uncertainty of the signal strength.

9.3 Systematic uncertainties

Systematic uncertainties are all the uncertainties that are not directly caused by limited sample sizes. In contrast to statistical uncertainties, they cannot be reduced by increasing the amount of data, but instead by improving the knowledge about the experiment, and of the signal and background modelling. The systematic uncertainties affect the Shape (S) and/or Normalization (N) of signal and background in the final discriminant distribution. The systematic uncertainties can be treated and reduced individually and uncorrelated from each other through the use of Nuisance Parameters (NPs) in the profile likelihood approach that was explained in Section 4.3. The different sources of systematic uncertainty that are considered in the analysis are presented in Table 9.11 and explained in brief in the following Sections. Table 9.11 also indicates whether the individual uncertainty affects only normalization, or shape and normalization.

Systematic uncertainty	Type	Components
Luminosity	N	1
Reconstructed Objects		
Electron trigger+reco+ID+isolation	SN	4
Electron energy scale+resolution	SN	2
Muon trigger+reco+ID+isolation	SN	8
Muon momentum scale+resolution+saggita	SN	5
Pileup modelling	SN	1
Jet vertex tagger	SN	1
Jet energy scale	SN	20
Jet energy resolution	SN	2
Missing transverse energy scale+resolution	SN	3
b-tagging efficiency	SN	5
c-mistag rate	SN	4
Light-mistag rate	SN	14
Mistag extrapolation $c \to \tau$	SN	1
Background and Signal Model		
$t\bar{t}$ cross section	N	1
$t\bar{t}+ \geq$1b: NLO Shape	SN	9
$t\bar{t}+ \geq$1c: NLO Shape	NS	1
$t\bar{t}+ \geq$1b: normalization	N (free floating)	1
$t\bar{t}+ \geq$1c: normalization	N (free floating)	1
$t\bar{t}+ \geq$3b: normalization	N	1
$t\bar{t}$ modelling: residual Radiation	NS	3
$t\bar{t}$ modelling: residual NLO generator	SN	3
$t\bar{t}$ modelling: residual parton shower+hadronisation	SN	3
W+jets normalisation	N	3
W+jets cross-section	N	1
Z+jets cross-section	N	1
Single top cross section	N	1
Single top model	SN	2
Diboson normalisation	N	1
Fakes normalization	SN	6
$t\bar{t}V$ cross section	N	4
$t\bar{t}V$ modelling	SN	2
tZ cross section	N	2
tWZ cross section	N	1
$t\bar{t}WW$ cross section	N	2
4-tops cross-section	N	1
Multijet normalisation	N	1
$t\bar{t}H$ cross section	N	2
$t\bar{t}H$ branching ratios	SN	1
$t\bar{t}H$ modelling	SN	1

Table 9.11: List of systematic uncertainties related to the object definitions. "SN" means that the uncertainty is taken on both shape and normalization, whereas "'N" means that the uncertainty is taken as normalization-only for all processes and channels affected. Some of the systematic uncertainties are split into several components for a more accurate treatment.

9.3.1 Luminosity

The 2016 integrated luminosity of the 13 TeV dataset (32.9 fb^{-1}) is measured with an uncertainty of 2.2% [87]. This systematic uncertainty is applied to all simulated samples where the event yield is derived from the MC . The multijet background is not affected by this uncertainty because. Whilst it is an MC sample, it is normalized to data using a minimum-bias triggered data sample.

9.3.2 Object reconstruction uncertainties

This section describes uncertainties associated with the object reconstruction and calibration, arising in the definitions of leptons, jets, heavy and light flavor tagging and E_T^{miss}.

Leptons

There are several sources of the uncertainty associated with the calibration of leptons arising from the trigger, identification, reconstruction, isolation and lepton momentum scale and resolution. The uncertainties affect the shape as well as the normalization. In general, the uncertainties associated with ith electrons (muons) include five (eight) components which have a small impact on the analysis.

Jets and missing transverse momentum

As is mentioned in Section 5.4, the main jet calibration applied are *jet energy scale* (JES), jet energy resolution (JER), and a cut on the *jet vertex fraction* (JVF). Associated uncertainties arise from the measurements of their efficiency. The JES measurements introduce some of the largest jet uncertainties. The overall uncertainty is approximately 5.5% for jets with $p_T > 100$ GeV [8]. The JES components are factorized into multiple independent sources (JES nuisance parameters):

- Pileup treatment uncertainties that are accounted for mismodelling in the MC simulation of the number of reconstructed, offset muons and primary vertices N_{pv}, pileup density ρ and pileup p_T term.
- Three uncertainties associated with the η interpolation (intercalibration).
- High-p_T jets: in-situ techniques derived from the single-particle response.

- One uncertainty associated with the punch-through correction.
- The 75 parameters from the in situ analyses combined to form 8 NPs. The uncertainties associated with in-situ techniques divided into different categories (statistical, detector, modeling, mixed) according to their origin e.g. γ/Z+jets balance.
- Two flavour-related uncertainties, to account for the fact that the response of the calorimeter and simulated jet composition differs for jets initiated by quarks and jets initiated by gluons.
- Another source of uncertainty is considered for jets originating from b-quarks which is measured by looking for the difference between an inclusive jet sample and a sample enriched in jets from b-quarks [138].

A systematic uncertainty associated to the JVF selection is estimated by using different generators for the MC simulation of $Z \rightarrow \mu^+\mu^-$ events and by changing the residual contamination from pileup jets.

Other jet-related uncertainties include the error on JER which is the uncertainty of width of Gaussian shape of the distribution of energy measurements for jets with the same true energy. The JER in data and MC are estimated from in-situ measurements as a function of jet η and p_T [139, 140].

The systematic uncertainties applicable to the E_T^{miss} come from the main-source uncertainties affecting high-p_T reconstructed objects which are directly translated into the missing transverse energy computation through equation 5.2. Uncertainties affecting the soft-jet and associated with leptons and jet energy scales and resolutions are smaller.

The systematic uncertainties applicable to the E_T^{miss} come from the soft-p_T term in equation 5.2. associated with leptons and jet energy scales and resolutions (track-based), which are propagated to E_T^{miss} and thus are included under the corresponding per-object uncertainty category.

Heavy and light Flavor tagging

The efficiency of the measured b, c and light-tagging differ in data and simulation. In order to take this difference into account, the b-tagging algorithms need to be calibrated in data. To that end, many methods have been developed to measure the b- and c-jet efficiency and the light-jet rejection rate in data. Therefore, the correction

scale factors (SFs=$\epsilon_{data}/\epsilon_{MC}$) are derived to describe the efficiency results. Systematic uncertainties arise from the derived correction SFs. Hence the SFs are derived as a function of p_T for jets originating from b and c-quarks. The light-jet rejection SFs depend on p_T and η. The systematic uncertainties are assumed to be uncorrelated between b, c and light jets. In order to propagate the multiple systematic uncertainties, which affect the the scale factors, into the NPs in a manageable way, a reduction in terms of 23 independent nuisance parameters through a diagonalisation method is used. The systematic uncertainties related to the b-tagging calibration are described with a total of five eigenvectors. The same method is performed to derive four eigenvectors on the c-tagging calibration and fourteen on light-jets (mistag) calibration. An additional uncertainty is included due to the extrapolation of the b- and c-tagging scale factors for jets with p_T beyond the kinematic reach of the data calibration samples used.

9.4 Modelling uncertainties

9.4.1 $t\bar{t}$+jets modelling

A number of systematic uncertainties affecting the modelling of the $t\bar{t}$+jets background are considered, summarized in Table 9.12 and described below.

The inclusive $t\bar{t}$ NNLO+NNLL production cross-section is assumed to have an uncertainty of $\pm 6\%$ [141] which results from the varying the input parameters in the generator, factorisation and renormalisation scales, the PDF, α_s, and the top quark mass.

An uncertainty associated with the choice of NLO generator is derived by comparing two alternative predictions, Powheg-Box+Pythia 8 and another generated with Sherpa with 5FS PDFs and a b-filter. Both samples are processed with the AF2 fast simulation [87], and propagating the resulting fractional difference to the nominal Powheg-Box+Pythia 8 full simulation prediction.

An uncertainty due to the choice of parton shower and hadronisation model is derived by comparing events produced by Powheg-Box interfaced to Pythia 8 or Herwig 7.

An uncertainty on modelling initial and final state radiation (ISR/FSR) is assessed with two alternative Powheg+Pythia 8 samples: one with the *hdamp* parameter set to $2m_t$, the renormalisation and factorisation scales set to half the nominal value, the A14 eigentune parameter Var3c increased and using the P2012 radHi

UE tune, giving radiation (referred to as "radHi"). The other one with the P2012 radLo UE tune, *hdamp* set to m_t, the renormalisation and factorisation scales set to double the nominal value and the A14 Var3c parameter decreased [142], giving less radiation (referred to as "radLow"). The uncertainties associated with the modelling of ISR/FSR are obtained by comparing the two samples "radHi" and "radLow" with the nominal PohegBox+Pythia 8 sample.

The above uncertainties are taken as uncorrelated between the $t\bar{t}$+light-jets, $t\bar{t}+\geq$ 1c and $t\bar{t}+\geq$1b processes because the diagrams of their production are different and the uncertainties may affect these processes in different regions in different ways, except for the uncertainty on the inclusive $t\bar{t}$ cross section.

In the case of $t\bar{t}+\geq$1b uncertainties on the SherpaOL NLO prediction, are considered. All alternative samples described above are reweighted to the NLO SherpaOL prediction in the same way as the nominal, in order to evaluate the relevant uncertainty. This uncertainty is considered as normalization-only systematic error. Different scale variations are considered to evaluate these uncertainties described in Table 9.13. Additionally, two alternative PDF sets (MSTW and NNPDF) and an alternative shower-recoil model scheme (CSS KIN) are considered. The uncertainty from multiple parton interactions (MPI) is assumed to be 50% based on studies of different underlying event tunes. In addition a 50% uncertainty on the normalization of the $t\bar{t}+\geq$3b component is introduced to account for the differences predicted by Powheg+Pythia 8 and SherpaOL.

Systematic source	How evaluated	$t\bar{t}$ categories
$t\bar{t}$ cross-section	$\pm$6%	All, correlated
NLO generator (*residual*)	Powheg-Box+Pythia 8 vs. Sherpa 5FS with b-filter	All, uncorrelated
Radiation (*residual*)	Variations of μ_R, μ_F, *hdamp* and the A14 Var3c parameters	All, uncorrelated
PS & hadronisation (*residual*)	Powheg-Box + Pythia 8 vs. Powheg-Box +Herwig 7	All, uncorrelated
$t\bar{t}+b\bar{b}$ renorm. scale (*reweighing*)	Up or down a by factor of two	$t\bar{t}+\geq$1b
$t\bar{t}+b\bar{b}$ resumm. scale (*reweighing*)	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t}+\geq$1b
$t\bar{t}+b\bar{b}$ global scales (*reweighing*)	set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t}+\geq$1b
$t\bar{t}+b\bar{b}$ shower recoil (*reweighing*)	Alternative model scheme	$t\bar{t}+\geq$1b
$t\bar{t}+b\bar{b}$ PDF (*reweighing*)	CT10 vs. MSTW or NNPDF	$t\bar{t}+\geq$1b
$t\bar{t}+\geq$3b reweighting	Up or down by 50%	$t\bar{t}+\geq$3b
$t\bar{t}+b\bar{b}$ MPI	Up or down by 50%	$t\bar{t}+\geq$1b

Table 9.12: List of the systematic uncertainties on the $t\bar{t}$+jets modelling. For the $t\bar{t}+\geq$1b background, the inclusive $t\bar{t}$ sample is reweighted to a NLO $t\bar{t}+b\bar{b}$ prediction; uncertainties on the inclusive sample are labelled *residual*, while those on the NLO prediction are labelled *reweighting*.

Scale	default	first variation	second variation (glosoft)	third variation (Q-CMMPS)
normalization μ_R	μ_{CMMPS}	$2\mu_{\text{CMMPS}}$, $\mu_{\text{CMMPS}}/2$	μ_{CMMPS}	μ_{CMMPS}
factorization μ_F	$H_{T,t}/2$	$H_{T,t}/2$	μ_{CMMPS}	$H_{T,t}/2$
resummation μ_Q	$H_{T,t}/2$	$H_{T,t}/2$	μ_{CMMPS}	μ_{CMMPS}

Table 9.13: Different scale variations considered in the NLO prediction for $t\bar{t}+ \geq$1b from SherpaOL.

9.4.2 Single-top modelling

The total cross section uncertainty for single-top-quark production is +5% −4% which is considered as a weighted average of the theoretical uncertainties on t-, Wt- and s-channel production [143–145].

Additional uncertainties associated with the modelling of ISR/FSR for the t- and Wt-channels are assessed by comparing the Powheg-Pythia 6 of the nominal samples with alternative samples where generator parameters have been varied (i.e. "radHi" and "radLow"). Additionally, for the same channels, an uncertainty due to the choice of parton shower and hadronisation model is derived by comparing events produced by Powheg-Box interfaced to Pythia or Herwig++. As is mentioned in the $t\bar{t}$+jets modelling in Section 9.4.1, alternative samples were generated using fast simulation. Comparisons are made with the Powheg-Box+Pythia sample using fast simulation, and then applied to the nominal sample, which was instead generated with full simulation.

The previous uncertainties are treated as fully correlated among single-top-quark production processes, but uncorrelated with the corresponding uncertainty on the $t\bar{t}$+jet samples which gives rise to an additional systematic uncertainty on the interference between Wt and $t\bar{t}$ production at NLO [146]. It is assessed by comparing the nominal sample, which uses the so called "diagram removal" scheme, with an alternative sample using the "diagram subtraction" scheme.

9.4.3 $t\bar{t}$+V modelling

The uncertainty on the $t\bar{t}$+V NLO cross-section prediction is established at 15% [147]. An uncertainty associated with the choice of NLO generator is derived by comparing two genrators; Powheg-Box Pythia 8 fast simulation and the MG5_aMC@NLO.

9.4.4 Signal modelling

The uncertainty on the $t\bar{t}H$ signal cross-section is $+10\%$ -13%, including contributions from scale and PDF uncertainties, which are treated as uncorrelated [148–152]. Additional uncertainties on the Higgs boson branching ratio to $b\bar{b}$ [130], WW and other final states are also included. The choice of parton shower and hadronization model uncertainty is accounted for by comparing with an alternative simulated sample generated with different showering and hadronization.

9.5 Fit results

A fit of all signal and background MC simulations to data collected in 2016 (corresponding to 32.9 fb^{-1} $\pm$2.2%) is performed in the nine analysis regions under the signal-plus-background hypothesis using non-KLFitter variables and the combination of non-KLFitter+KLFitter variables which are listed in Table 9.8. The details of the adopted fit method have been discussed in Section 4.3. The abbreviated names of all the nuisance parameters used in the fit is explained in appendix D. The results of the fit to Asimov dataset[1]and data and limits on $t\bar{t}H(H \to b\bar{b})$ production are given in Sections 9.5.1 and 9.5.2, respectively.

9.5.1 Expected results

This Section discusses expected results for non-KLFitter and the combination of non-KLFitter+KLFitter MC samples obtained by performing fits to the Asimov dataset which is constructed from the nominal Monte Carlo expectation and contains no statistical fluctuations, see [153]. By using the Asimov data set one can easily obtain the median significance. The actual data will contain statistical fluctuations and thus the observed significance and the upper limit are not in general equal to the median and this leads to higher or weaker signal strength μ than is obtained by the Asimov dataset.

The distributions of the fitted nuisance parameters corresponding to theoretical and instrumental systematic uncertainties for KLFitter and non-KLFitter samples are shown in Figures 9.11 and E.1, respectively. The mean best fit values (the black dots) for each parameter which are called *pulls* are zero by construction when the Asimov dataset is used and this refers to the ideal case of perfect Monte Carlo modeling of the signal and backgrounds. The size of the constraints (the uncertainties on each pull) on the nuisance parameters in the Asimov dataset are a statement about the power of the data to constrain the parameters. If a nuisance parameter is constrained in the Asimov dataset, it is reasonable that it will also be constrained in the fit to data. That appears in particular for those tight constraints associated with large uncertainties on $t\bar{t}$ modeling, since the analysis is dominated by this background. The corresponding NP correlation matrix for KLFitter and non-KLFitter samples (only NPs with a correlation coefficient of at least 20% with any other parameter are

[1]The Asimov dataset is defined as the one where the pseudodata (artificial data) is equal to the expectation value. It simply combines the background and signal MC events already used in the analysis.

displayed) can be found in Figures 9.12 and E.2, respectively. In general, there is no difference between KLFitter and non-KLFitter NP distributions and correlations.

The expected uncertainties on the normalization factors for $t\bar{t}+ \geq$HF and (Heavy Flavour), as well as the signal strength uncertainty are slightly lower for KLFitter than the non-KLFitter factors, which indicates that the KLFitter variables improve the measurements on these factors, see Figures 9.13 and E.3. Table 9.15 shows that the KLFitter variables make some improvement, and that the expected limit is lower for KLFitter than non-KLFitter when using the Asimov dataset.

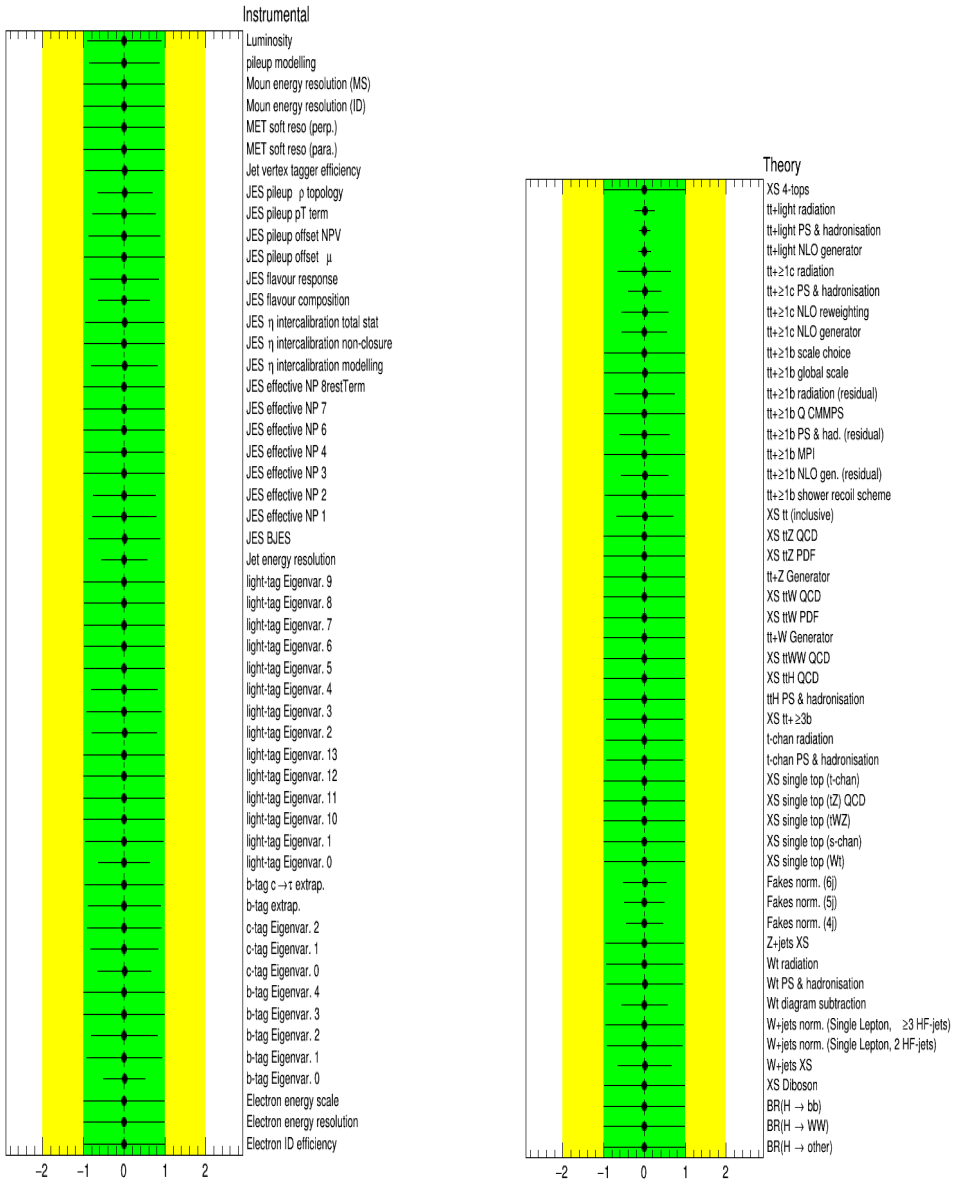


Figure 9.11: KLFitter nuisance parameters corresponding to the Instrumental (left plot) and Theory (right plot) systematic uncertainties in fits to the Asimov dataset.

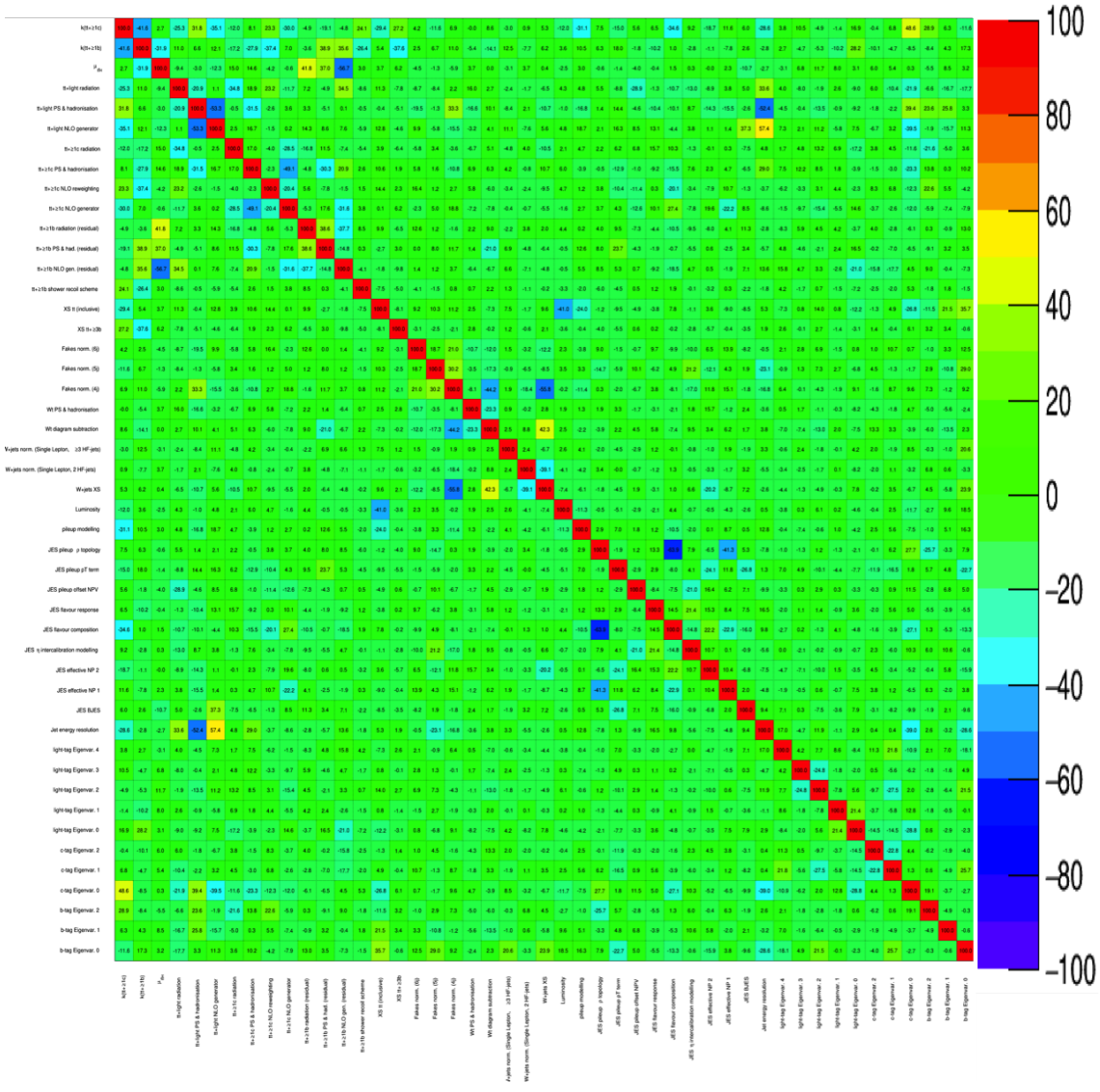


Figure 9.12: KLFitter correlation matrix between NPs corresponding to the fit to Asimov dataset.

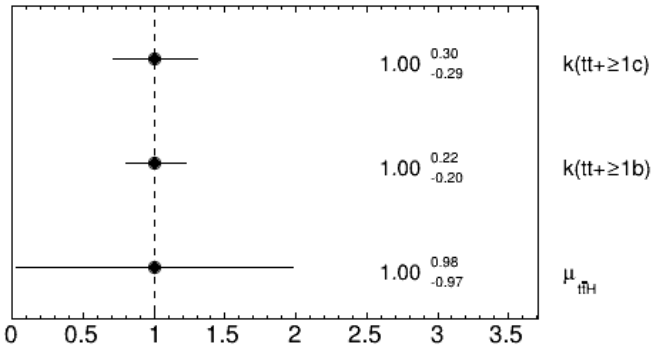


Figure 9.13: KLFitter normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ and Parameter of interest (signal strength μ) components corresponding to the fit to Asimov dataset.

9.5.2 Fits to data

A fit to the data in the nine analysis regions is performed under the signal-plus-background hypothesis, and the fitted NPs for the KLFitter and non-KLFitter samples are shown in Figures 9.14 and E.4, respectively. The corresponding correlation matrix for the fitted NPs for both KLFitter and non-KLFitter (only NPs with a correlation coefficient of at least 20% with any other parameter are displayed) can be found in Figures 9.16 and E.6, respectively. Figure 9.19 shows the distribution of pre- and post-fit events per region for KLFitter samples. It is clear that the fit improves the data/MC ratio in all regions. The relative uncertainties are also much smaller in the post-fit plot. The pre-fit and post-fit yields for data and MC simulation processes in both control and signal regions for KLFitter and non-KLFitter samples (which are shown in the previous Figures) are reported in the Tables 9.16, 9.17, E.1 and E.2, respectively. The fake normalization uncertainty is set to $\pm 50\%$ in each of the background regions. This background contribution is at low BDT weight values and hence is set to zero in the signal regions.

The most relevant pulls and constraints that are associated with the considered regions are discussed in the following:

- In general, the observed constraints on the systematic uncertainties are compatible with the expected values from the Asimov fit.
- The pulls and constraints show similar results for KLFitter and non-KLFitter samples.
- As for the Asimov fit in Figure 9.11, the $t\bar{t}$ modeling systematic parameters present an uncertainty smaller than 1σ. This is a result of the large initial value of the systematic error with respect to the data statistical error. The fit then chooses to change the central value and the size of the systematic uncertainty to be smaller than the original one.
- Some systematic parameters have been pulled by the fit to give a best fit to data. An example is the strong negative pull on $t\bar{t}+ \geq 1c$ radiation. This result comes from the use of the nominal and two other samples to provide two working points corresponding to increased and decreased radiation. The fit value has chosen a strong negative pull (around -2.5 σ).
- Other systematic uncertainties are not discussed, since their pulls and constraints do not significantly affect the sensitivity of the analysis.

- Correlation matrices show that there are no major issues, with most correlations <10% and all correlations < 62%. That means the variables used are satisfactorily independent except some correlations between $t\bar{t}$ modelling systematic uncertainties in the top left corner of Figure 9.16 which may explain the high constraints of these systematic uncertainties.

The fitted values of the two free-floating normalization factors and the signal strength for KLFitter and non-KLFitter are displayed in Figures 9.17 and 9.18, respectively. In the KLFitter case, the fit chose the cross-section of $t\bar{t}+ \geq 1b$ to be almost equal to theory, and to be higher by 35% than the theory for the $t\bar{t}+ \geq 1c$ cross-section. The KLFitter results show improvement for the $k(t\bar{t}+ \geq 1c)$ fitted value $(1.35^{+0.27}_{-0.25})$ compared to the non-KLFitter fitted value $(1.5^{+0.29}_{-0.28})$.

As for the signal strength μ, in general, its value is around 2σ away from what is expected in the Asimov dataset. A discrepancy with respect to the SM prediction is apparent in the KLFitter analysis. In order to understand the origin of this difference, a comparison of the non-KLFitter analysis was made with respect to the ICHEP 2016 semi-leptonic analysis [61], as summarised in Table 9.14. An upward fluctuation was also seen in the analysis. However, the increase in luminosity, a different b-tag working point (WP) and a change of the default top MC have made this difference more significant. The additional of KLFitter variables increased the central value further, whilst reducing the uncertainties, leading to the observed overall 3 σ difference. The difference between the central value of signal strength in the Asimov set and measured data may be explained by the following: a statistical fluctuation in data (i.e. more data is needed) or there are some systematic uncertainties which were not assessed, or were incorrectly modeled.

The subsequent published analysis [154] redefined the approach with respect to signal binning and used continuous b-tagging in order to improve the sensitivity and reduce the systematic uncertainties. This reduced the size of the discrepancy such that the results were consistent with the SM. However, a complete re-analysis on this basis, including the KLFitter variables, is beyond the scope of this thesis.

Systematic type		KLFitter	non-KLFitter
significance fit to Data	Expected	1.18	1.12
	Observed	4.25	3.48
limit fit to Data	Expected	1.88	2.01
	Observed	5.40	4.90
limit fit to Asimov	Expected	1.95	2.11

Table 9.15: The KLFitter and non-KLFitter significance and limit ratio.

Analysis Differences	ICHEP-2016	non-KLFitter analysis
1-luminosity [fb^{-1}]	13	33
2-b-tagging WP [%]	70	77
3-top MC generator	Powheg-Box+Pythia 8	Powheg-Box+Pythia 6
signal strength μ	$1.6^{+1.1}_{-1.1}$	$3.12^{+1.03}_{-0.90}$

Table 9.14: Comparison of non-KLFitter analysis with the ICHEP-2016 analysis.

The expected significance goes from 1.12 σ to 1.18 σ when the KLFitter variables are used. The high value of observed significance in both the KLFitter and non-KLFitter samples indicates the possibility of the presence of a signal. When using data, the observed significance goes from 3.48 σ to 4.25 σ where the KLFitter variables are included, see Table 9.15.

Figure E.5 demonstrates the effect on the best-fit value μ with respect to the top 20 nuisance parameters and the constraints provided by the data. The largest effect arises from the $t\bar{t}+ \geq 1b$ modelling, and that is reasonable because it is the dominant background in the signal-rich regions. The $t\bar{t}H$ systematics (QCD, PDF, PS and hadronization) all have an important effect on the μ measurement.

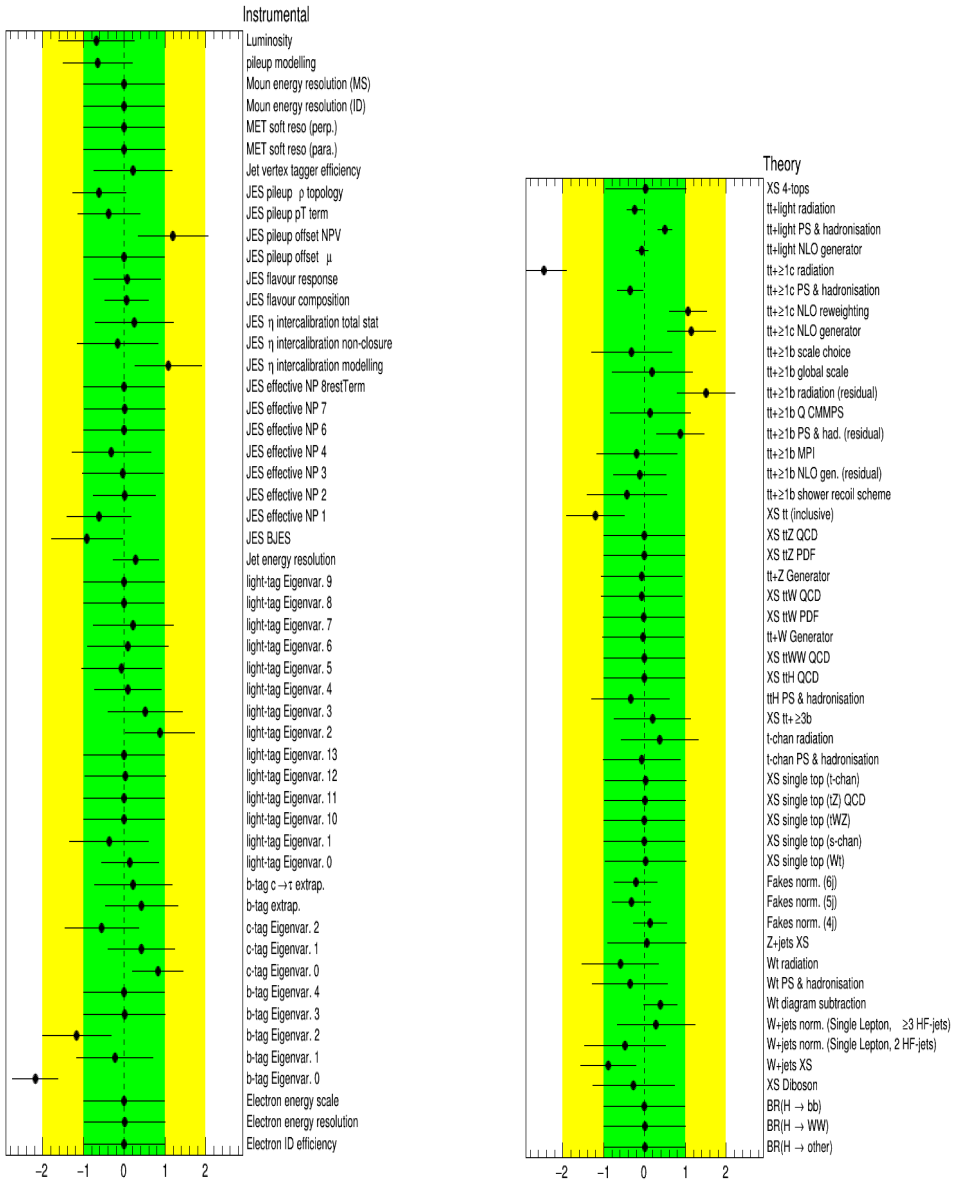


Figure 9.14: The KLFitter nuisance parameters corresponding to the Instrumental (left plot) and Theory (right plot) systematic uncertainties in fits to data.

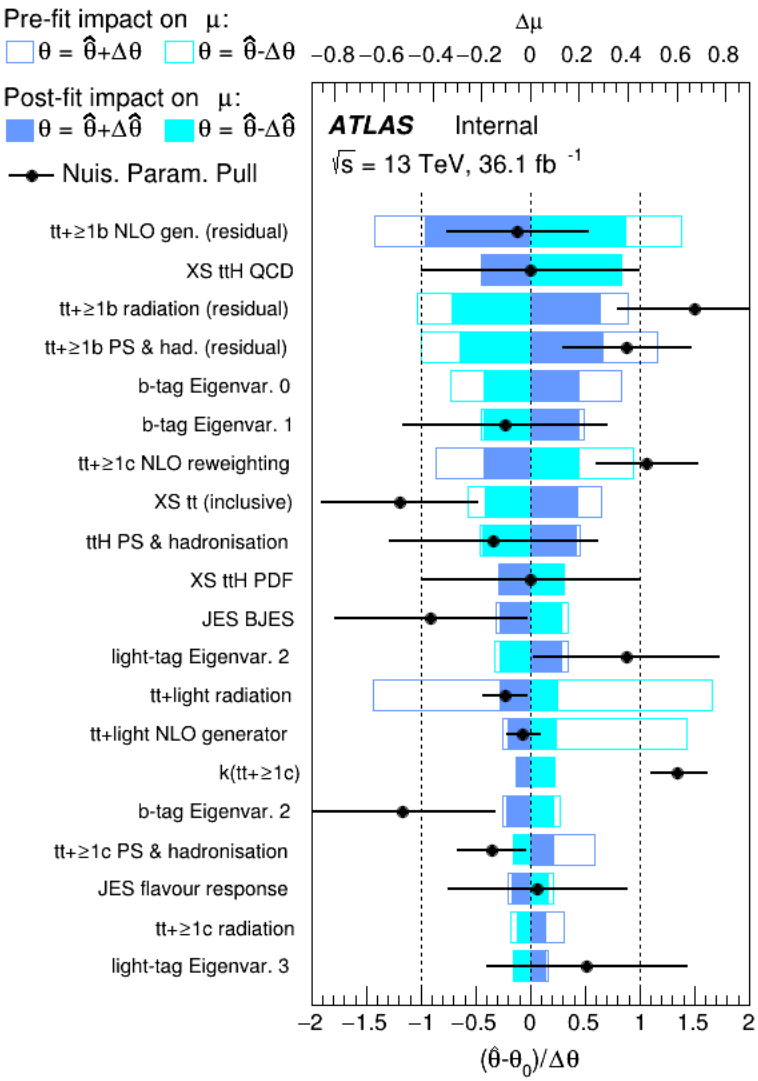


Figure 9.15: The fitted values of the KLFitter 20 NPs which have the largest effect on the measured signal strength corresponding to the fit to data.

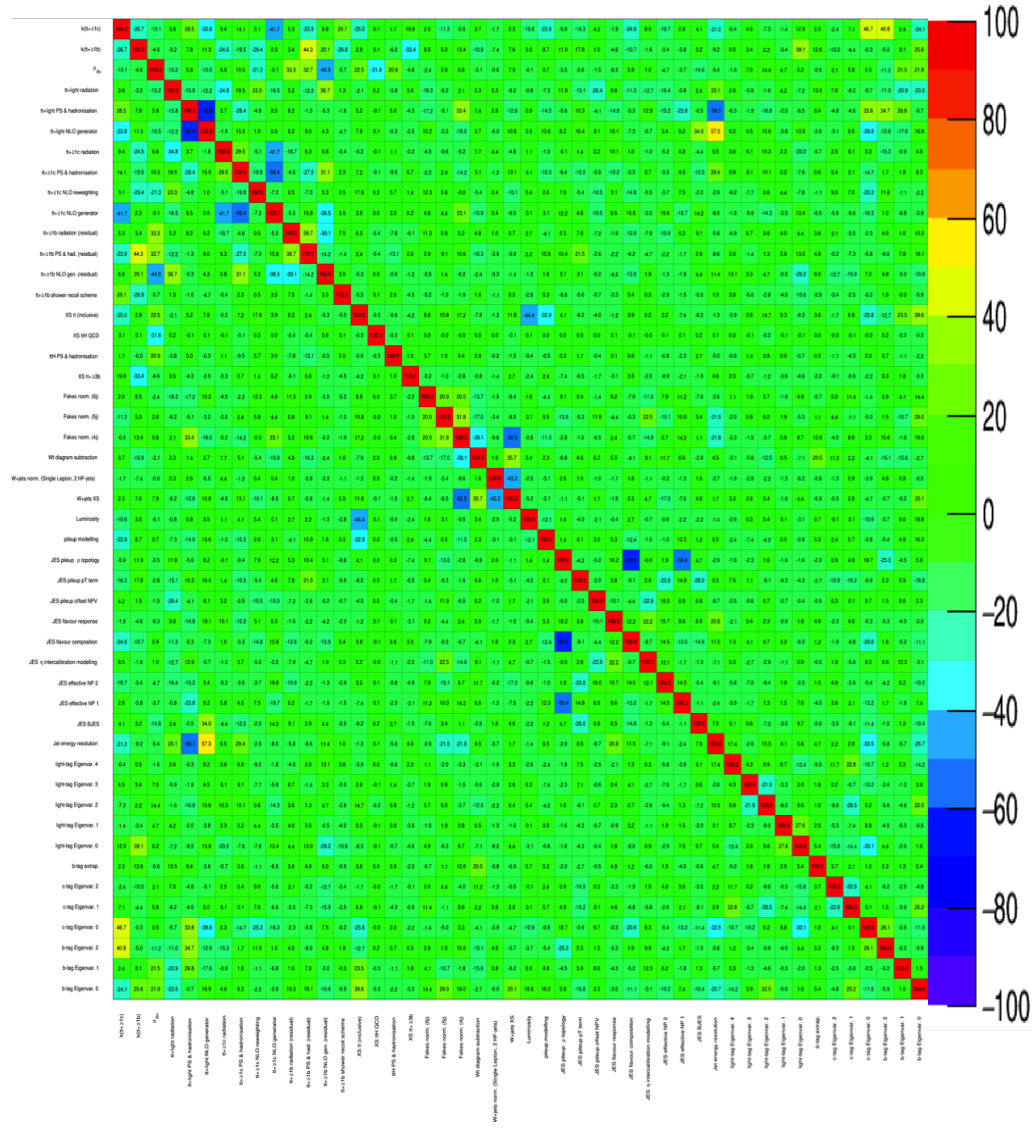


Figure 9.16: The KLFitter correlation matrix between NPs corresponding to fit to data.

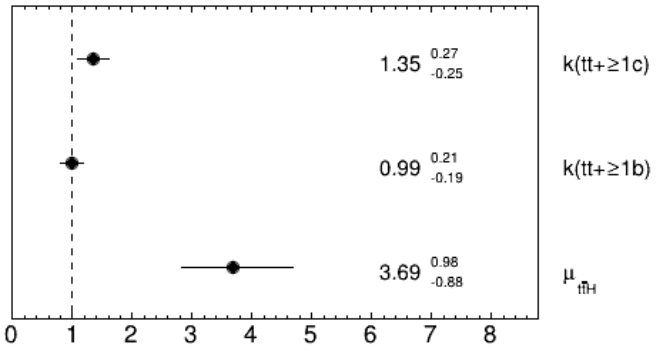


Figure 9.17: The KLFitter normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ and Parameter of interest (signal strength μ) components corresponding to the fit to data.

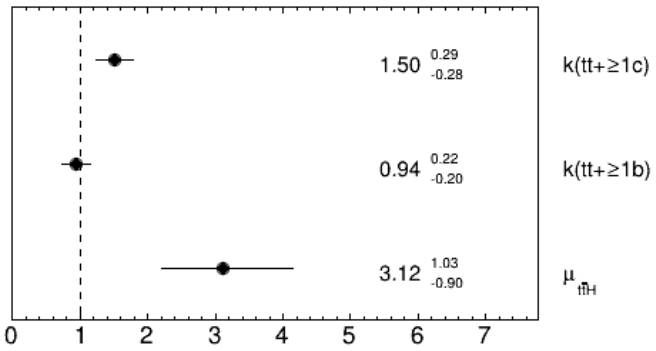


Figure 9.18: The non-KLFitter normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ and Parameter of interest (signal strength μ) components corresponding to the fit to data.

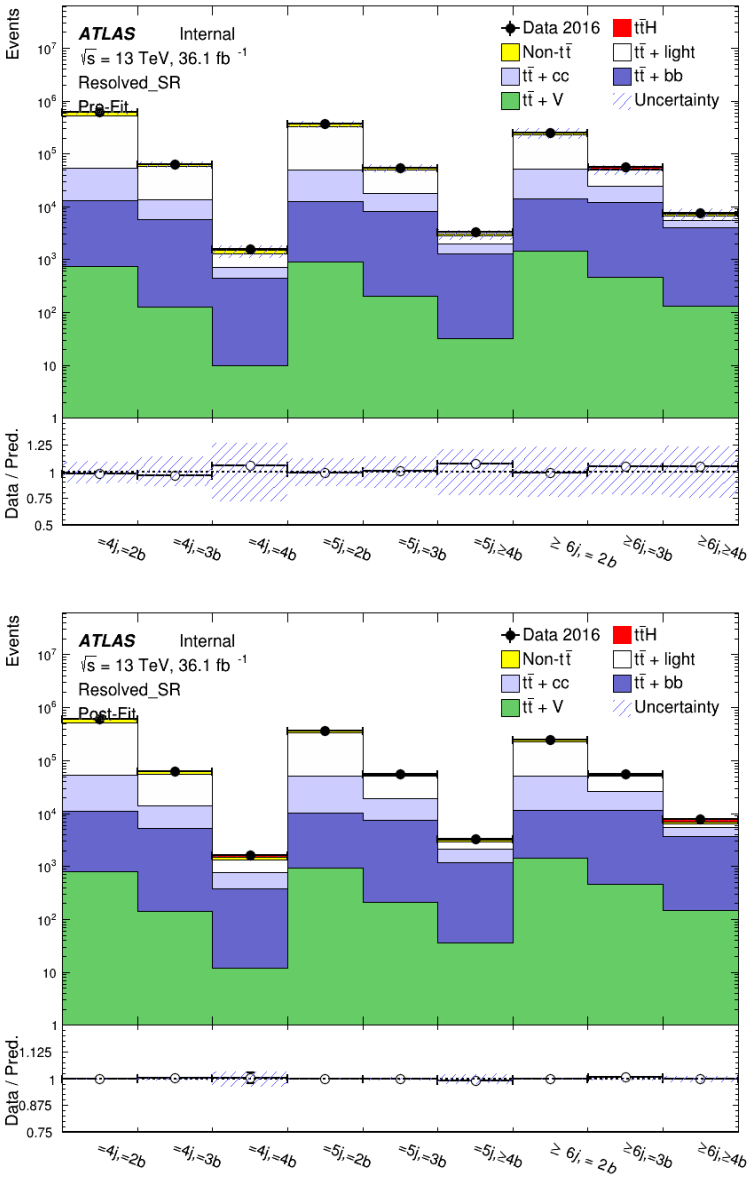


Figure 9.19: Comparison between the data and KLFitter MC prediction per-region yield pre-fit (top) and post-fit (bottom), fit to data.

	= 4j, = 2b	= 4j, = 3b	= 4j, = 4b	= 5j, = 2b	= 5j, = 3b	= 5j, ≥ 4b	≥ 6j, = 2b	≥ 6j, = 3b	≥ 6j, ≥ 4b
Fakes	22 800 ± 11 400	1870 ± 944	76.0 ± 41.8	8730 ± 4370	914 ± 466	0 ± 0	5080 ± 2550	0 ± 0	0 ± 0
$t\bar{t}$ + light	475 000 ± 49 900	43 800 ± 6320	546 ± 293	279 000 ± 41 900	32 000 ± 5900	873 ± 302	176 000 ± 51 200	25 000 ± 7550	1140 ± 681
$t\bar{t}$ + ≥1c	40 900 ± 11 800	8070 ± 2480	275 ± 75.8	36 800 ± 6750	9590 ± 1770	684 ± 240	37 300 ± 9890	12 400 ± 3560	1520 ± 649
$t\bar{t}$ + ≥1b	12 100 ± 2720	5640 ± 1860	443 ± 178	11 500 ± 1780	7890 ± 1810	1270 ± 349	12 600 ± 1920	11 500 ± 2170	3950 ± 1000
$t\bar{t} + V$	107 ± 10.7	29.1 ± 4.07	2.73 ± 0.524	73.1 ± 8.19	19.0 ± 2.40	3.94 ± 0.787	92.1 ± 14.2	44.6 ± 12.1	21.4 ± 8.50
Single top tchan	11 800 ± 1870	882 ± 144	38.5 ± 26.3	3960 ± 782	460 ± 97.9	35.2 ± 11.0	1730 ± 437	309 ± 94.5	48.8 ± 27.1
Single top schan	905 ± 116	67.3 ± 11.1	2.80 ± 0.869	295 ± 47.5	32.4 ± 5.39	1.88 ± 0.677	114 ± 25.6	20.0 ± 4.43	2.71 ± 1.11
Single top Wtchan	19 900 ± 4680	1920 ± 569	58.5 ± 40.3	11 000 ± 3090	1560 ± 510	81.6 ± 35.1	6470 ± 2420	1320 ± 562	164 ± 91.0
Diboson	1120 ± 573	83.0 ± 43.4	4.36 ± 2.55	617 ± 320	64.6 ± 34.5	5.30 ± 3.14	530 ± 282	82.0 ± 43.1	8.62 ± 5.53
$t\bar{t} + W$	303 ± 59.7	34.6 ± 8.44	1.04 ± 0.577	402 ± 69.2	65.4 ± 12.7	3.65 ± 1.24	603 ± 105	145 ± 29.2	18.4 ± 4.56
$t\bar{t} + Z$	332 ± 54.6	64.9 ± 13.5	6.22 ± 1.96	432 ± 57.0	118 ± 18.7	25.0 ± 4.94	761 ± 113	275 ± 42.9	94.7 ± 17.1
W + jets	31 100 ± 14 700	1930 ± 996	34.1 ± 19.1	13 800 ± 6520	1260 ± 624	55.3 ± 29.3	8770 ± 4260	1150 ± 597	121 ± 64.6
Z + jets	5470 ± 2100	367 ± 170	8.36 ± 3.93	2420 ± 964	272 ± 152	11.6 ± 5.23	1570 ± 673	207 ± 87.4	22.2 ± 9.58
$t\bar{t}H$ ($H \to b\bar{b}$)	101 ± 12.8	71.8 ± 11.0	12.9 ± 2.56	133 ± 14.8	138 ± 17.7	48.0 ± 8.86	210 ± 29.3	292 ± 34.8	183 ± 33.7
$t\bar{t}H$ ($H \to WW$)	30.4 ± 5.10	3.66 ± 0.855	0.135 ± 0.144	55.3 ± 6.63	9.05 ± 1.39	0.478 ± 0.269	157 ± 20.2	42.2 ± 6.18	5.69 ± 1.27
$t\bar{t}H$ ($H \to$ other)	31.6 ± 4.63	4.60 ± 0.785	0.273 ± 0.0823	54.3 ± 7.68	10.7 ± 1.71	1.17 ± 0.470	122 ± 17.0	36.9 ± 5.66	6.68 ± 1.30
Total	622 000 ± 62 700	64 800 ± 8840	1510 ± 415	370 000 ± 49 200	54 400 ± 8190	3100 ± 665	252 000 ± 58 800	52 800 ± 11 200	7310 ± 1750
Data 2016	609458	62421	1590	363928	54678	3339	248492	55102	7650

Table 9.16: Pre-fit event yields for the KLFitter single lepton channel regions fit to data. The three signal regions are highlighted in bold text.

	= 4j, = 2b	= 4j, = 3b	= 4j, = 4b	= 5j, = 2b	= 5j, = 3b	= 5j, ≥ 4b	≥ 6j, = 2b	≥ 6j, = 3b	≥ 6j, ≥ 4b
Fakes	24 400 ± 4730	2000 ± 405	81.2 ± 23.5	7190 ± 2200	753 ± 246	0 ± 0	4480 ± 1410	0 ± 0	0 ± 0
$t\bar{t}$ + light	473 000 ± 9450	41 100 ± 1890	541 ± 91.2	278 000 ± 7120	30 900 ± 1780	805 ± 132	177 000 ± 5750	24 500 ± 1750	1140 ± 208
$t\bar{t}$ + ≥1c	41 900 ± 9320	9040 ± 1820	385 ± 67.8	41 100 ± 7390	11 900 ± 1910	962 ± 154	39 000 ± 6210	15 100 ± 2150	1670 ± 306
$t\bar{t}$ + ≥1b	10 200 ± 2200	5030 ± 1100	373 ± 78.7	9330 ± 1940	7280 ± 1390	1150 ± 189	10 100 ± 2050	10 900 ± 1850	3550 ± 434
$t\bar{t} + V$	110 ± 9.44	32.8 ± 3.16	3.50 ± 0.443	73.8 ± 6.38	20.1 ± 1.87	4.74 ± 0.579	88.5 ± 10.9	44.1 ± 11.2	24.2 ± 9.10
Single top tchan	12 200 ± 1370	938 ± 88.7	43.4 ± 24.8	4120 ± 496	530 ± 74.8	44.3 ± 9.56	1820 ± 258	347 ± 74.8	57.7 ± 24.3
Single top schan	949 ± 67.3	74.0 ± 6.90	3.00 ± 0.696	307 ± 22.9	34.3 ± 3.35	2.38 ± 0.584	115 ± 9.80	20.5 ± 2.41	3.26 ± 0.730
Single top Wtchan	20 900 ± 2950	1990 ± 350	47.1 ± 20.2	10 600 ± 1420	1490 ± 254	91.4 ± 25.7	5350 ± 948	1060 ± 243	139 ± 43.3
Diboson	980 ± 537	76.2 ± 42.2	4.23 ± 2.53	496 ± 273	55.5 ± 31.1	4.43 ± 2.75	436 ± 240	68.9 ± 38.2	7.24 ± 4.56
$t\bar{t} + W$	328 ± 58.5	35.9 ± 7.86	1.03 ± 0.546	421 ± 68.0	66.8 ± 11.5	3.65 ± 1.06	606 ± 87.5	143 ± 23.9	18.4 ± 3.46
$t\bar{t} + Z$	356 ± 50.4	71.6 ± 12.5	7.32 ± 1.91	447 ± 54.4	124 ± 17.1	28.3 ± 4.24	749 ± 89.5	271 ± 33.4	104 ± 13.8
W + jets	18 100 ± 5980	1230 ± 543	23.2 ± 11.4	8070 ± 2700	770 ± 295	45.1 ± 19.4	5010 ± 1690	776 ± 324	85.6 ± 36.4
Z + jets	5630 ± 1920	371 ± 146	9.78 ± 3.85	2460 ± 845	239 ± 126	13.9 ± 5.36	1550 ± 545	210 ± 73.4	24.4 ± 8.93
$t\bar{t}H$ ($H \to b\bar{b}$)	375 ± 86.9	311 ± 68.0	60.7 ± 12.9	447 ± 108	566 ± 125	218 ± 46.2	607 ± 157	1060 ± 244	790 ± 170
$t\bar{t}H$ ($H \to WW$)	123 ± 28.4	14.9 ± 3.42	0.674 ± 0.564	221 ± 50.4	34.5 ± 8.64	1.46 ± 0.945	588 ± 134	156 ± 35.6	21.6 ± 5.02
$t\bar{t}H$ ($H \to$ other)	129 ± 29.5	18.4 ± 4.19	1.02 ± 0.305	220 ± 50.2	41.5 ± 10.4	4.56 ± 1.57	455 ± 105	135 ± 31.4	25.8 ± 5.98
Total	609 000 ± 1280	62 400 ± 457	1590 ± 59.1	364 000 ± 978	54 800 ± 339	3380 ± 77.7	248 000 ± 721	54 800 ± 346	7660 ± 122
Data 2016	609458	62421	1590	363928	54678	3339	248492	55102	7650

Table 9.17: Post-fit event yields for the KLFitter single lepton channel regions fit to data. The three signal regions are highlighted in bold text.

9.6 Summary

A comparison was made between KLFitter and non-KLFitter variables in terms of a MultiVariate Analysis by providing four BDTs to illustrate the improvement in sensitivity of the analysis when the KLFitter variables are used. The treatment of several sources of systematic uncertainties was described. The fit results show an improvement in terms of the higher value of signal strength μ. A lower value of $k(t\bar{t}+ \geq 1c)$ normalization factor and also lower associated uncertainties are obtained. The lower value of the expected limit also derives from the efficiency of using KLFitter variables. The use of KLFitter variables thus enables a signal to be established more readily for a greater data sample size. The observed signal strength obtained in the $t\bar{t}H$ channel is:

$$\mu = 3.69^{+0.98}_{-0.88}$$

The result is outside the expected range and somewhat above the SM expectation. It is interpreted as a statistical upward fluctuation i.e. a null result with an upper limit of $\mu < 5.4$ (95% CL).

Chapter 10

Conclusion

In this thesis, searches for new phenomena involving the production of a Higgs boson in association with a top quark pair, with the decay $H \to b\bar{b}$ and the single-lepton channel of the $t\bar{t}$ pair, were presented. The analysis uses the LHC proton-proton collision dataset recorded by the ATLAS detector in 2016 at $\sqrt{s}$ =13 TeV, corresponding to an integrated luminosity of 32.9 fb^{-1}. The general analysis strategy employs a kinematic fit to the signal-rich regions using the KLFitter package and constructs variables for the analysis that are directly sensitive to how signal-like an events is. Later on, these variables are combined with the ranked variables constructed by an earlier analysis, refereed to as the non-KLFitter study. The discrimination power of the variables is combined into one single variable by using a MultiVariate Analysis (BDTs) as part of the TMVA package.

The implementation of KLFitter is studied by testing several KLFitter jet selection modes due to the apparently low Higgs-finding efficiency when using only the 6 jets selection mode (kB6). The result of this study leads to the conclusion that the 8 jet selection mode (kB8) has the highest Higgs boson reconstruction efficiency by providing an extra two slots for the b-tags in the hope that two bs quarks from the Higgs or top decays will be included. However, the TMVA study led to the exclusion of the kB7 and kB8 jet selection modes and adopted the kB6 in the rest of the analysis.

The measurements of the $t\bar{t}H$ signal strength and its upper limit at a centre-of-mass energy of pp collision of 13 TeV were presented. The $t\bar{t}H$ signal strength and its upper limit have been measured. A significant excess of events above the background expectation was found thus the observed upper limit (significance) was high, 5.4 (4.25 σ) which indicates the presence of signal which is hard to exclude. The

KLFitter expected limit was lower compared to non-KLFitter one which refer to the improvement in the analysis sensitivity when the KLFitter variables are used. The ratio of the measured $t\bar{t}H$ signal cross section to the Standard Model expectation is found to be $\mu = 3.69^{+0.98}_{-0.88}$. This signal strength result is a reflection of the statistical fluctuation based on a limited dataset. A signal strength larger than 5.4 can be excluded at the 95% confidence level:

$$\mu < 5.4 \text{ (95\% CL)}.$$

Appendices

Appendix A

Derivation of $p_{\nu z}$ from the W mass constraint

In order to derive how to calculate $p_{\nu z}$ from the W mass constraint, we start by writing down the four momentum of the charged hard lepton p_l and the neutrino p_ν that the addition of these momenta in the same event gives the parent leptonically-decaying W boson particle.

$$p_l = (p_{lx}, p_{ly}, p_{lz}, E_l), \tag{A.1}$$

$$p_\nu = (p_{\nu x}, p_{\nu y}, p_{\nu z}, E_\nu), \tag{A.2}$$

then when neglecting the masses of both the neutrino and the charged lepton as their masses are small in comparison to their momenta, the invariant mass of W boson is

$$m_W^2 = (E_l + E_\nu)^2 - (p_{lx} + p_{\nu x})^2 - (p_{ly} + p_{\nu y})^2 - (p_{lz} + p_{\nu z})^2 \tag{A.3}$$

Therefore, according to the previous approximation:

$$E_l = |p_l|, \tag{A.4}$$

$$E_\nu = |p_\nu| \tag{A.5}$$

Equation A.3 can then be rearranged to give:

$$2E_l E_\nu = m_W^2 + 2p_{lx}p_{\nu x} + 2p_{ly}p_{\nu y} + 2p_{lz}p_{\nu z} \tag{A.6}$$

Squaring equation A.6 gives:

$$4(p_{lx}^2 + p_{ly}^2 + p_{lz}^2)(p_{\nu x}^2 + p_{\nu y}^2 + p_{\nu z}^2) = (m_W^2 + 2p_{lx}p_{\nu x} + 2p_{ly}p_{\nu y} + 2p_{lz}p_{\nu z})^2 \quad \text{(A.7)}$$

Putting everything to one side and sorting in powers of $p_{\nu z}$,

$$\begin{aligned} 4(p_{lx}^2 + p_{ly}^2)\mathbf{p}_{\nu z}^2 - (4m_W^2 p_{lz} + 8p_{lx}p_{\nu x}p_{lz} + 8p_{ly}p_{\nu y}p_{lz})\mathbf{p}_{\nu z} \\ + (4p_{ly}^2 p_{\nu x}^2 + 4p_{lz}^2 p_{\nu x}^2 + 4p_{lx}^2 p_{\nu y}^2 + 4p_{lz}^2 p_{\nu y}^2 \\ - 8p_{lx}p_{\nu x}p_{ly}p_{\nu y} - 4p_{lx}p_{\nu x}m_W^2 - 4p_{ly}p_{\nu y}m_W^2 - m_W^4) = 0. \end{aligned} \quad \text{(A.8)}$$

The quadratic equation formula for obtaining of $p_{\nu z}$ is thus:

$$p_{\nu z} = (-b \pm \sqrt{b^2 - 4ac})/2a \quad \text{(A.9)}$$

The solution of the previous equation has two cases:

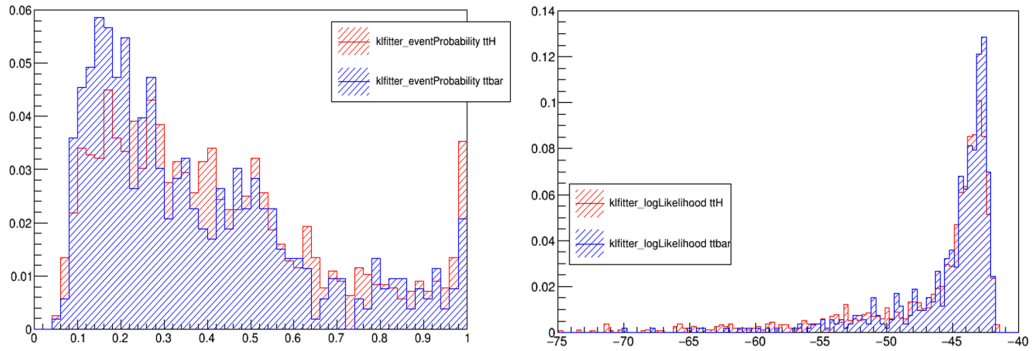
- case 1: if $b^2 - 4ac \geq 0$, two real distinct roots are available; $p_{\nu z0}$ and $p_{\nu z1}$.
- case 2: if $b^2 - 4ac < 0$, the the assumption $b^2 - 4ac = 0$ is used, which leads to the third solution $p_{\nu z2} = -b/2a$.

Then, the neutrino energy is calculated according to the equation:

$$E_\nu^2 = E_{\nu x}^2 + E_{\nu y}^2 + p_{\nu z}^2 \quad \text{(A.10)}$$

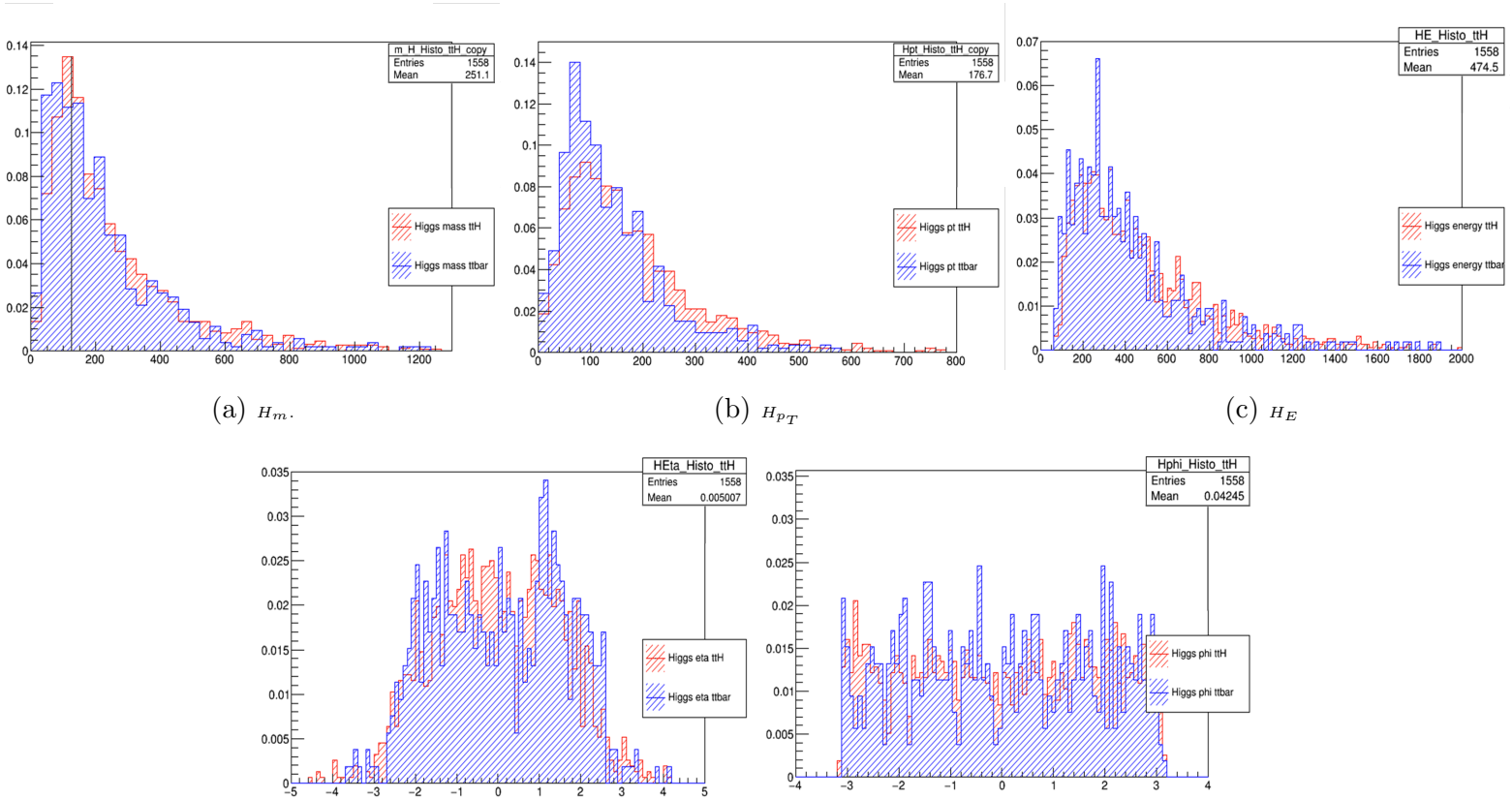
Appendix B

Figures of Study Stage I



(a) klf_eventProbability (b) klf_logLikelihood

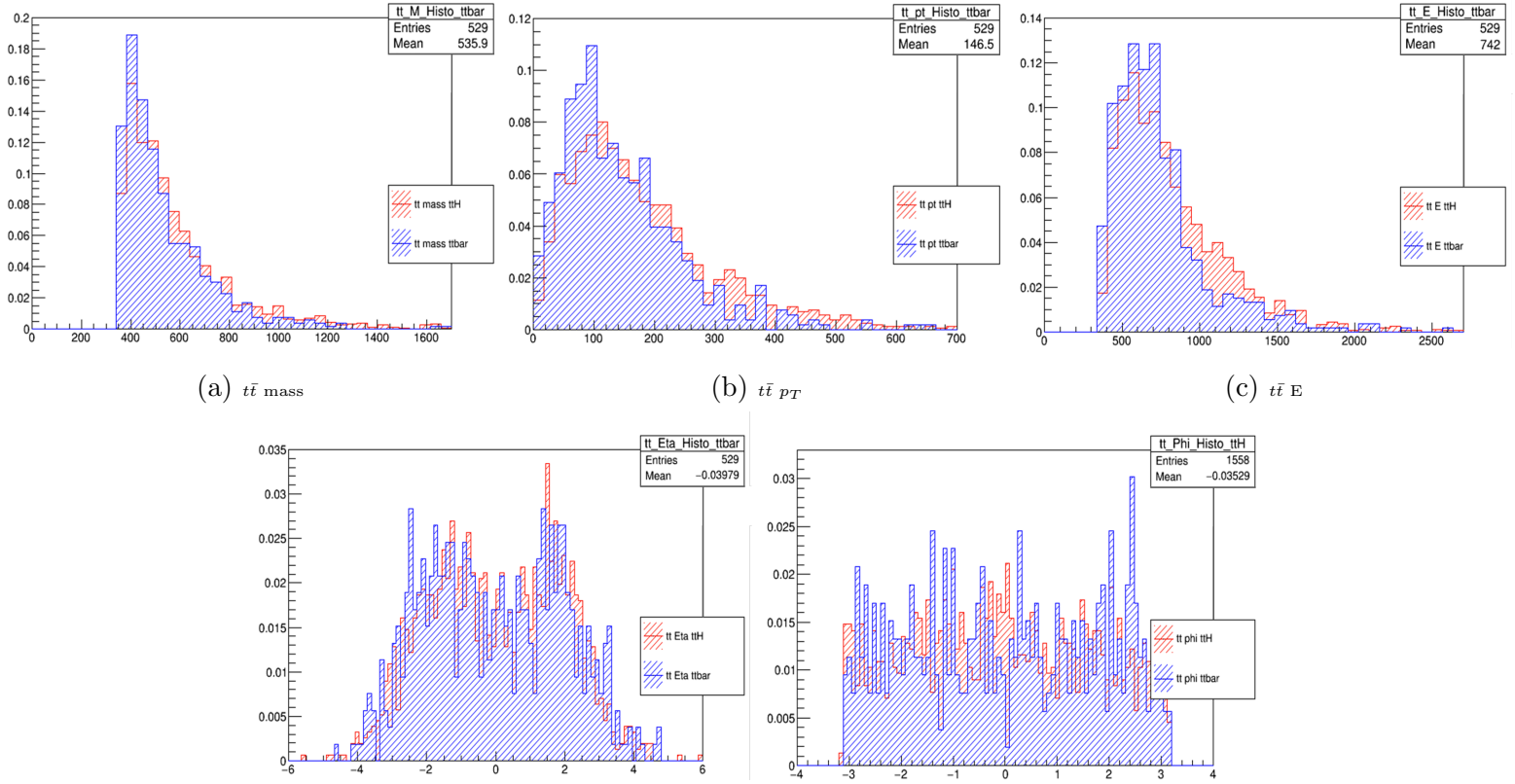
Figure B.1: Normalised distributions showing the shapes of klftter_eventProbability (left) and klfitter_logLikelihood (right) for $t\bar{t}H$ signal (red) and $t\bar{t}$ +jets samples (blue), during the first stage of the study (i.e. [6j, 2b] inclusive region, kBtagPrioritySixJets mode).



(a) H_m. (b) H_{p_T} (c) H_E

(d) H_η (e) H_Φ

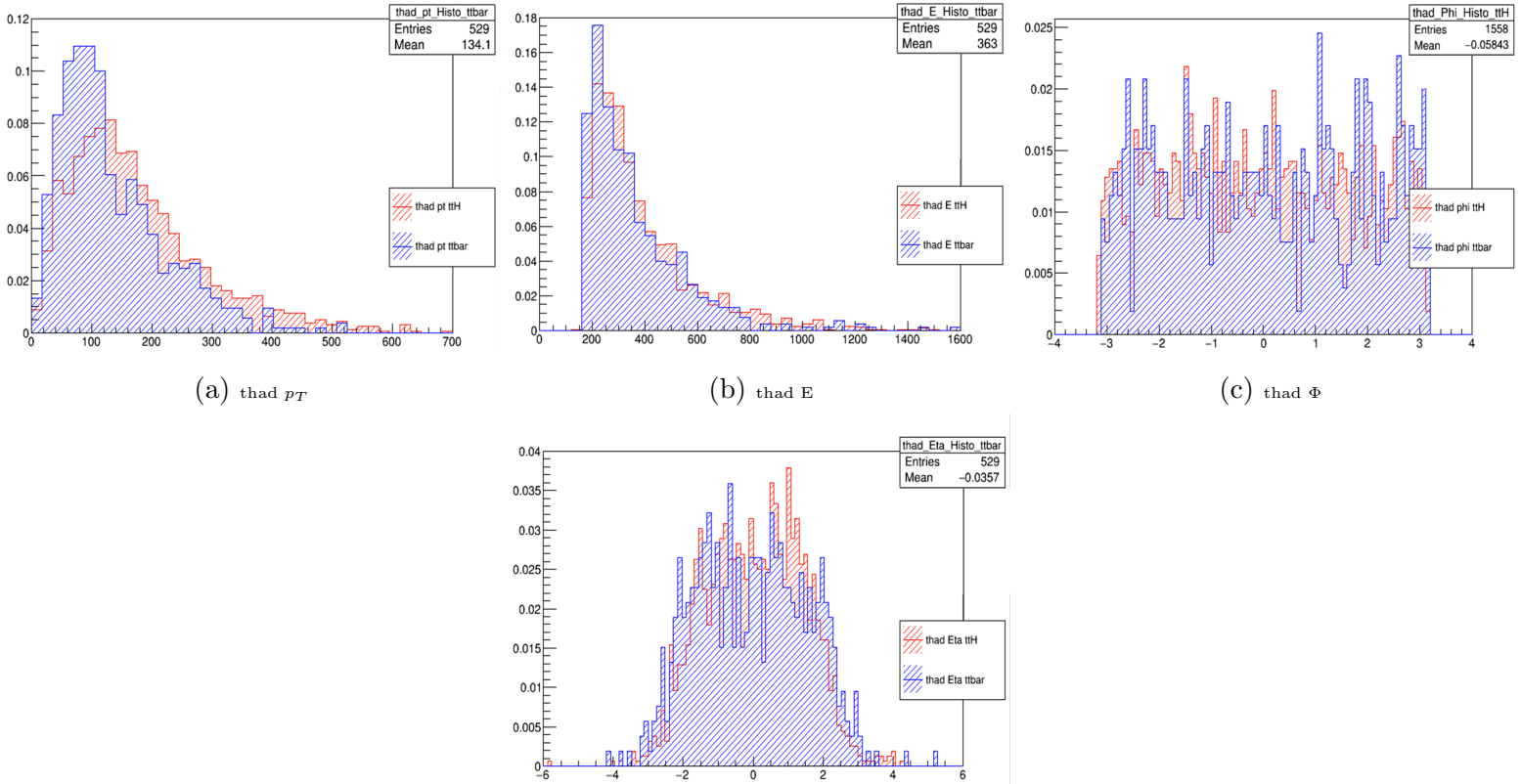
Figure B.2: Normalised distributions showing the shapes of reconstructed Higgs boson mass, p_T, Energy, η and Φ for $t\bar{t}H$ signal (red) and $t\bar{t}$ +jets samples (blue), during the first stage of the study (i.e. [6j, 2b] inclusive region, kBtagPrioritySixJets mode).



(a) $t\bar{t}$ mass (b) $t\bar{t}$ p_T (c) $t\bar{t}$ E

(d) $t\bar{t}$ η (e) $t\bar{t}$ Φ

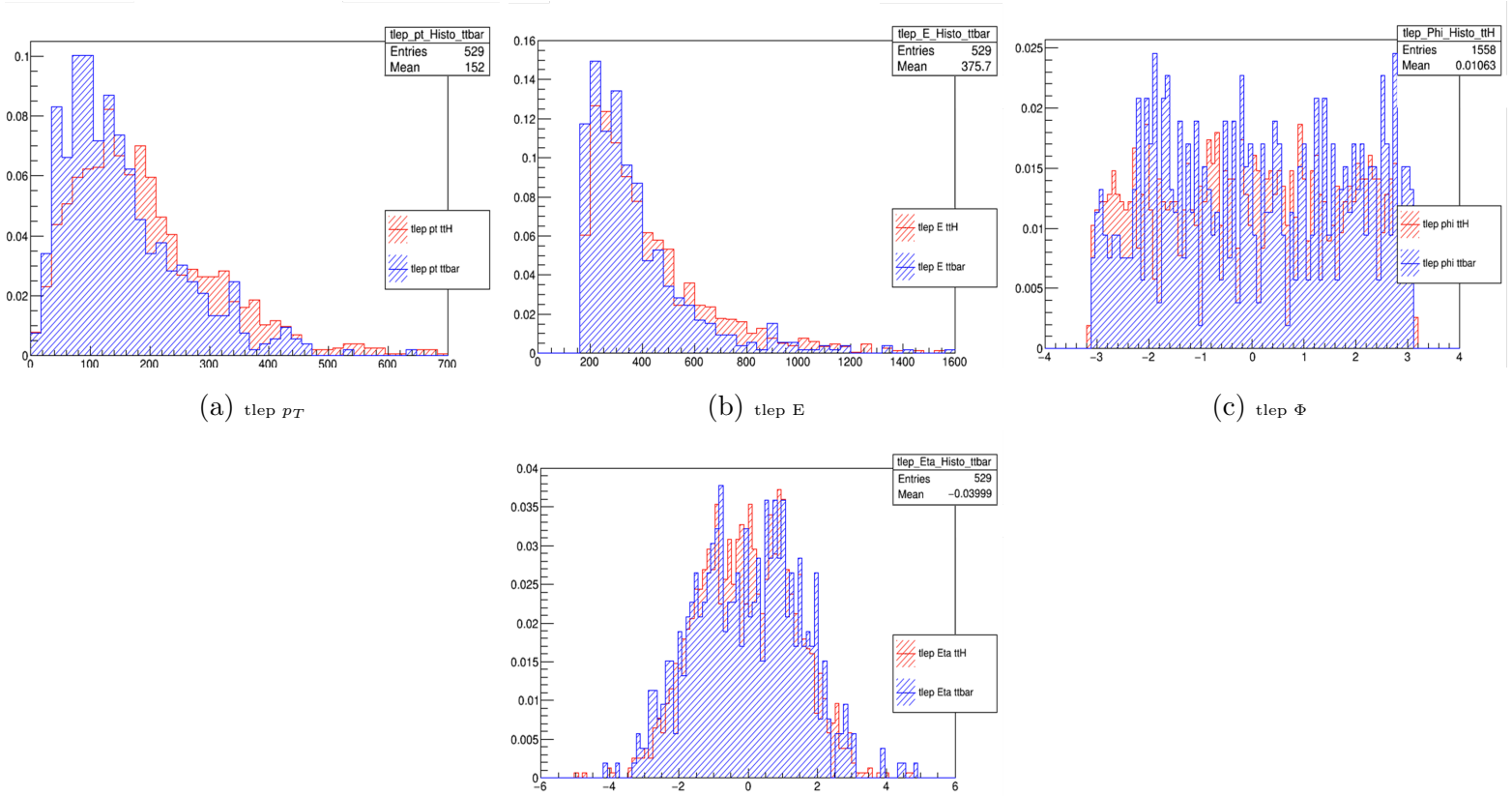
Figure B.3: Normalised distributions showing the shapes of reconstructed $t\bar{t}$ system mass, p_T, Energy, Eta and Phi for $t\bar{t}H$ signal (red) and $t\bar{t}$ +jets samples (blue), during the first stage of the study (i.e. [6j, 2b] inclusive region, kBtagPrioritySixJets mode).



(a) thad p_T (b) thad E (c) thad Φ

(d)

Figure B.4: Normalised distributions showing the shapes of reconstructed thad p_T, Energy, η and Φ for $t\bar{t}H$ signal (red) and $t\bar{t}$ +jets samples (blue), during the first stage of the study (i.e. [6j, 2b] inclusive region, kBtagPrioritySixJets mode).



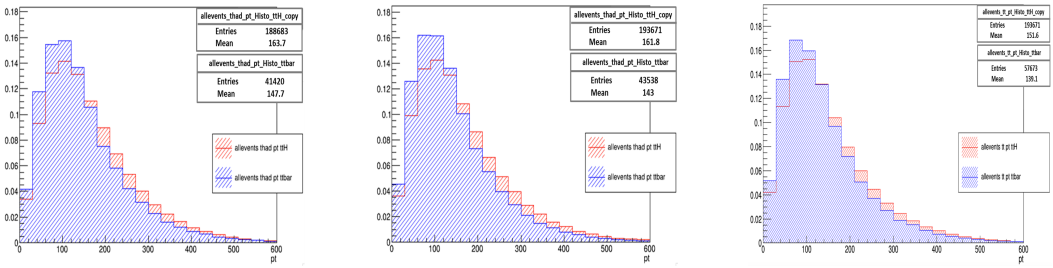
(a) tlep p_T (b) tlep E (c) tlep Φ

(d) tlep η

Figure B.5: Normalised distributions showing the shapes of reconstructed tlep p_T, Energy, η and Φ for $t\bar{t}H$ signal (red) and $t\bar{t}$ +jets samples (blue), during the first stage of the study (i.e. [6j, 2b] inclusive region, kBtagPrioritySixJets mode).

Appendix C

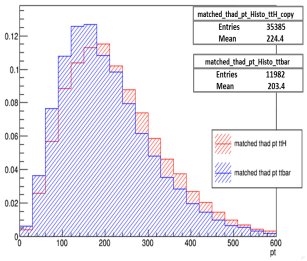
Figures of the weak variables in Study Stage III thad,tlep and $t\bar{t}$ system



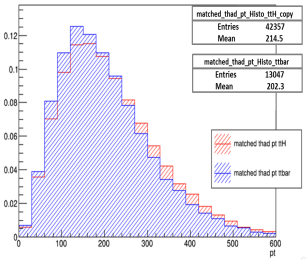
(a) $thad_{p_T}^{reco}$ kb6

(b) $thad_{p_T}^{reco}$ kb7

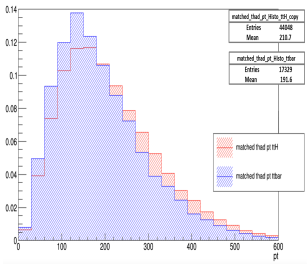
(c) $thad_{p_T}^{reco}$ kb8



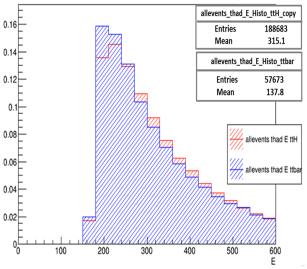
(d) $thad_{p_T}^{matched}$ kb6



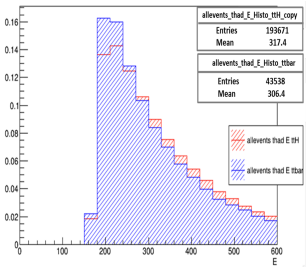
(e) $thad_{p_T}^{matched}$ kb7



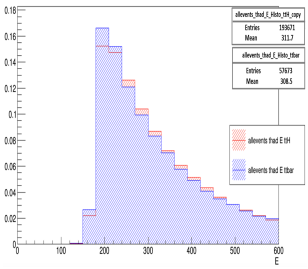
(f) $thad_{p_T}^{matched}$ kb8



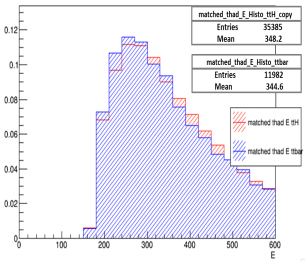
(g) $thad_{E}^{reco}$ kb6



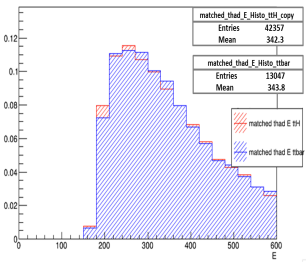
(h) $thad_{E}^{reco}$ kb7



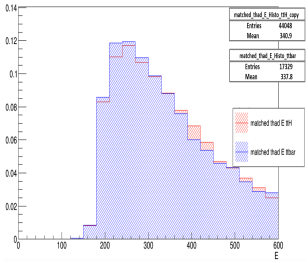
(i) $thad_{E}^{reco}$ kb8



(j) $thad_{E}^{matched}$ kb6

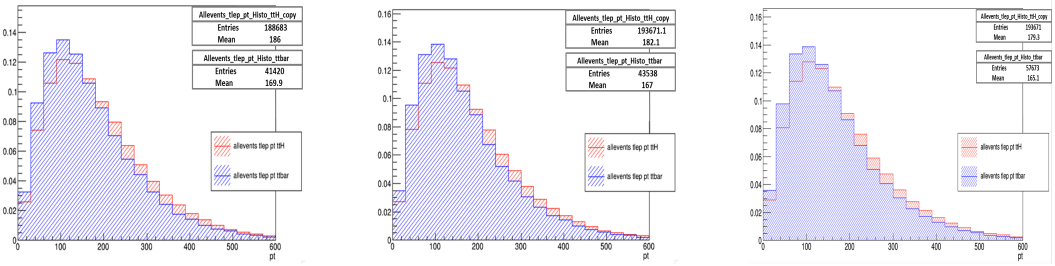


(k) $thad_{E}^{matched}$ kb7

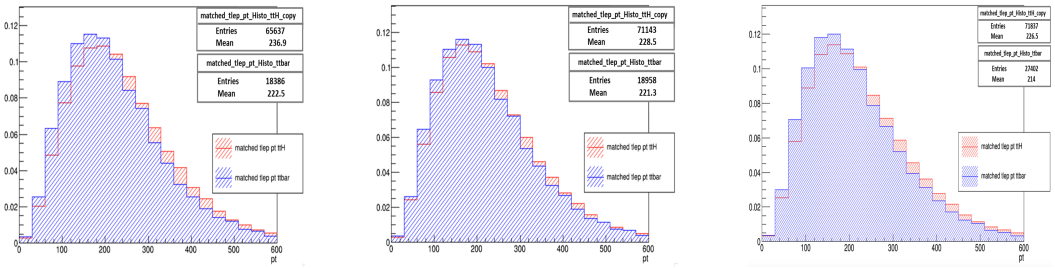


(l) $thad_{E}^{matched}$ kb8

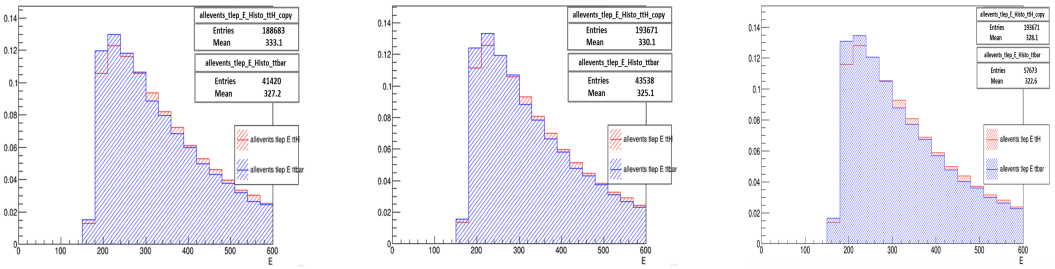
Figure C.1: kBtagPriority6j/7j/8j p_T and E of thad. Additional non-discreminating variables (not used in TMVA).



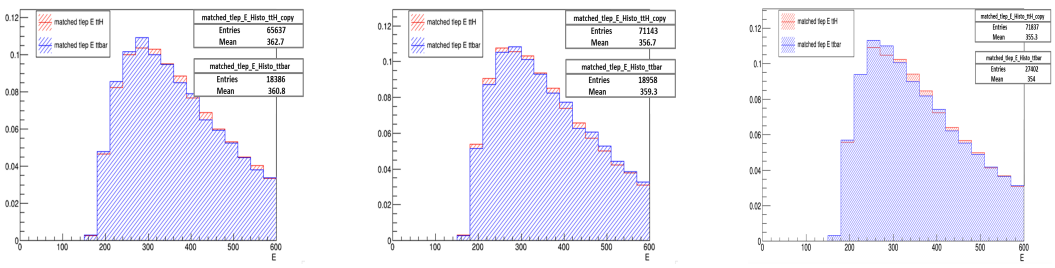
(a) $tlep_{pT}^{reco}$ kb6 (b) $tlep_{pT}^{reco}$ kb7 (c) $tlep_{pT}^{reco}$ kb8



(d) $tlep_{pT}^{matched}$ kb6 (e) $tlep_{pT}^{matched}$ kb7 (f) $tlep_{pT}^{matched}$ kb8

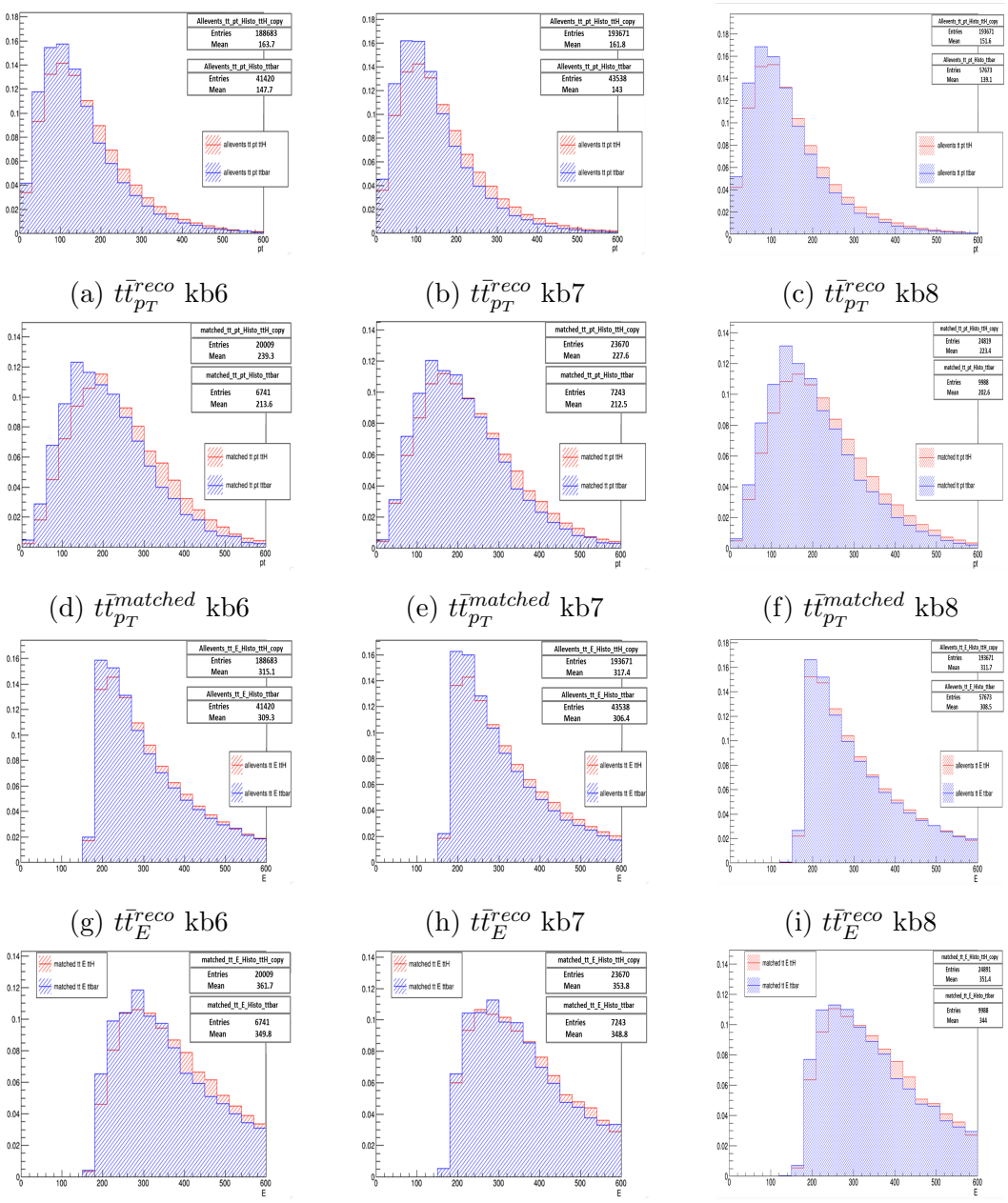


(g) $tlep_{E}^{reco}$ kb6 (h) $tlep_{E}^{reco}$ kb7 (i) $tlep_{E}^{reco}$ kb8



(j) $tlep_{E}^{matched}$ kb6 (k) $tlep_{E}^{matched}$ kb7 (l) $tlep_{E}^{matched}$ kb8

Figure C.2: kBtagPriority6j/7j/8j p_T and E of tlep. Additional non-discreminating variables (not used in TMVA).



(a) $t\bar{t}^{reco}_{p_T}$ kb6 (b) $t\bar{t}^{reco}_{p_T}$ kb7 (c) $t\bar{t}^{reco}_{p_T}$ kb8

(d) $t\bar{t}^{matched}_{p_T}$ kb6 (e) $t\bar{t}^{matched}_{p_T}$ kb7 (f) $t\bar{t}^{matched}_{p_T}$ kb8

(g) $t\bar{t}^{reco}_{E}$ kb6 (h) $t\bar{t}^{reco}_{E}$ kb7 (i) $t\bar{t}^{reco}_{E}$ kb8

(j) $t\bar{t}^{matched}_{E}$ kb6 (k) $t\bar{t}^{matched}_{E}$ kb7 (l) $t\bar{t}^{matched}_{E}$ kb8

Figure C.3: kBtagPriority6j/7j/8j p_T and E of $t\bar{t}$ system. Additional non-discreminating variables (not used in TMVA).

Appendix D

Glossary for SM Analysis Nuisance Parameter Plots

This appendix is a key to the abbreviated names of the nuisance parameters used in the fit output.

- Luminosity: Luminosity
- Pileup: Pile-up uncertainty or pileup modelling.
- Electron:
  - identification: Electron ID efficiency.
  - reconstruction: Electron reconstruction efficiency.
  - isolation: Electron isolation efficiency.
  - energy scale: Electron energy scale
  - energy resolution: Electron energy resolution.
  - trigger: Electron trigger efficiency.
- Muon:
  - identification: Muon ID efficiency (syst), Muon ID efficiency (stat).
  - reconstruction: Muon TTVA efficiency (syst), Muon TTVA efficiency (stat).
  - isolation: Muon isol efficiency (syst), Muon isol efficiency (stat).
  - resolution in the tracker: Moun energy resolution (ID).
  - resolution in the muon system: Moun energy resolution (MS).
  - trigger: Muon trig efficiency (syst), Muon trig efficiency (stat).
  - Muon sagitta: Muon sagitta residual biais, Muon sagitta ρ topology.

- Jet vertex tagger (JVT): Jet vertex tagger efficiency.
- Jet energy scale:
  - JES effective NP 1, JES effective NP 2, JES effective NP 3, JES effective NP 4, JES effective NP 5, JES effective NP 6, JES effective NP 7, JES effective NP 8restTerm.
  - JES η intercalibration: total stat, modelling and non-closure.
  - JES flavour composition, JES flavour response.
  - JES pileup offset μ, JES pileup offset NPV, JES pileup p_T term and JES pileup ρ topology.
  - Calorimeter response to b-jets: JES BJES.
  - Punch-through correction: JES punchthrough.
- Jet energy resolution: : JER or Jet energy resolution.
- Missing transverse momentum:
  - Soft term energy scale: MET soft scale
  - Soft term resolution: MET soft reso (perp.), MET soft reso (para.).
- b-tagging uncertainties:

  5 eigenvectors corresponding to b-jet p_T bins ordered from the largest to the smallest: b-tag Eigenvar. 0, b-tag Eigenvar. 1, b-tag Eigenvar. 2, b-tag Eigenvar. 3, b-tag Eigenvar. 4
- c-tagging uncertainties:

  4 eigenvectors corresponding to c-jet p_T bins ordered from the largest to the smallest: c-tag Eigenvar. 0, c-tag Eigenvar. 1, c-tag Eigenvar. 2, c-tag Eigenvar. 3
- light-tagging uncertainties:

  14 eigenvectors corresponding to 7 light-jet p_T bins and two jet η regions ordered from the largest to the smallest: l-tag Eigenvar. 0, l-tag Eigenvar. 1, l-tag Eigenvar. 2, l-tag Eigenvar. 3, l-tag Eigenvar. 4, l-tag Eigenvar. 5, l-tag Eigenvar. 6, l-tag Eigenvar. 7, l-tag Eigenvar. 8, l-tag Eigenvar. 9, l-tag Eigenvar. 10, l-tag Eigenvar. 11, l-tag Eigenvar. 12, l-tag Eigenvar. 13
- high-p_T extrapolation uncertainty for b- and c-tagging: b-tag c$\rightarrow \tau$ extrap., b-tag extrap.
- Free-floating normalisation factors:
  - $\mu_{t\bar{t}H}$: signal strength for $t\bar{t}H$, the parameter of interest in the fit.
  - normalisation of the $t\bar{t}+\geq 1b$ background component: $k(t\bar{t}+\geq 1b)$
  - normalisation of the $t\bar{t}+\geq 1c$ background component: $k(t\bar{t}+\geq 1c)$
- Signal modelling uncertainties:
  - signal cross-section uncertainty from scale variations: XS ttH QCD
  - signal cross-section uncertainty from PDF choice: XS ttH PDF
  - Parton Showering and hadronisation uncertainty: ttH PS & hadronisation

- uncertainty on the Higgs decay rate to $b\bar{b}$: BR(H $\rightarrow$ bb)
- uncertainty on the Higgs decay rate to WW: BR(H $\rightarrow$ WW)
- uncertainty on the Higgs decay rate to other final states: BR(H $\rightarrow$ other)

- $t\bar{t}$ background modelling uncertainties:
  - total $t\bar{t}$ cross-section uncertainty: "XS tt (inclusive)
  - uncertainty on the scale choice on $t\bar{t}+ \geq 1b$ reweighting, derived varying the default renormalisation scale by a factor of two up and down: $t\bar{t}+ \geq 1b$ scale choice
  - uncertainty on the scale choice on $t\bar{t}+ \geq 1b$ reweighting, derived using an alternative softer scale for both factorisation and resummation scales: $t\bar{t}+ \geq 1b$ global scale
  - uncertainty on the scale choice on $t\bar{t}+ \geq 1b$ reweighting, derived using an alternative softer scale only for the resummation scale: $t\bar{t}+ \geq 1b$ Q CMMPS
  - uncertainty from the shower recoil scheme e (CSS KIN) in the $t\bar{t}+ \geq 1b$ reweighting: $t\bar{t}+ \geq 1b$ shower recoil scheme
  - one of the two uncertainties from PDF choice on $t\bar{t}+ \geq 1b$ reweighting: $t\bar{t}+ \geq 1b$ MSTW PDF
  - one of the two uncertainties from PDF choice on $t\bar{t}+ \geq 1b$ reweighting: $t\bar{t}+ \geq 1b$ NNPDF PDF
  - uncertainty on the normalisation of the $t\bar{t}+ \geq 1b$
  - uncertainty on the modelling of the underliying event: $t\bar{t}+ \geq 1b$ UE modelling
  - $t\bar{t}+ \geq 1b$ Matrix Element model: $t\bar{t}+ \geq 1b$ NLO gen. (residual)
  - $t\bar{t}+ \geq 1b$ Radiation model: $t\bar{t}+ \geq 1b$ radiation (residual)
  - $t\bar{t}+ \geq 1b$ Parton shower model: $t\bar{t}+ \geq 1b$ PS & had. (residual)
  - $t\bar{t}+ \geq 1c$ NLO modelling: $t\bar{t}+ \geq 1c$ NLO reweighting
  - $t\bar{t}+ \geq 1c$ Matrix Element model: $t\bar{t}+ \geq 1c$ NLO generator
  - $t\bar{t}+ \geq 1c$ Radiation model: $t\bar{t}+ \geq 1c$ radiation
  - $t\bar{t}+ \geq 1c$ Parton shower model: $t\bar{t}+ \geq 1c$ PS & hadronisation
  - $t\bar{t}+ \geq$light Matrix Element model: $t\bar{t}+ \geq$light NLO generator
  - $t\bar{t}+ \geq$light Radiation model: $t\bar{t}+ \geq$light radiation
  - $t\bar{t}+ \geq$light Parton shower model: $t\bar{t}+ \geq$light PS & hadronisation

- Other small background modelling uncertainties:
  - inclusive cross-section uncertainty on the single-top Wt-channel process: XS single top (Wt)
  - inclusive normalisation uncertainty on the single-top non-W t-channel processes: XS single top (t- & s-chan)
  - the uncertainty on the interference between Wt and $t\bar{t}$ production at NLO: Wt diagram subtraction

- Wt Radiation model: Wt radiation
- Wt Parton shower model: Wt PS & hadronisation
- uncertainty on the fraction of W+HF jets in the single lepton regions with 2 b-tagged jets: W+jets norm. (Single Lepton, 2 HF-jets)
- uncertainty on the fraction of W+HF jets in the single lepton regions with 3 b-tagged jets: W+jets norm. (Single Lepton, ≥ 3 HF-jets)
- W+jets cross-section uncertainty :W+jets XS
- Z+jets cross-section uncertainty: Z+jets XS
- $t\bar{t} + W$ cross-section uncertainty from scale variations: XS ttW QCD
- $t\bar{t} + W$ cross-section uncertainty from PDF choice: XS ttW PDF
- $t\bar{t} + W$ generator: tt+W Generator
- $t\bar{t} + Z$ cross-section uncertainty from scale variations: XS ttZ QCD
- $t\bar{t} + Z$ cross-section uncertainty from PDF choice: XS ttZ PDF
- $t\bar{t} + Z$ generator: tt+Z Generator
- Diboson cross section uncertainty: XS Diboson
-
- Multijet uncertainty: Fakes norm. (4j), Fakes norm. (5j) and Fakes norm. (6j)

Appendix E

Non-KLFitter Plots and Tables

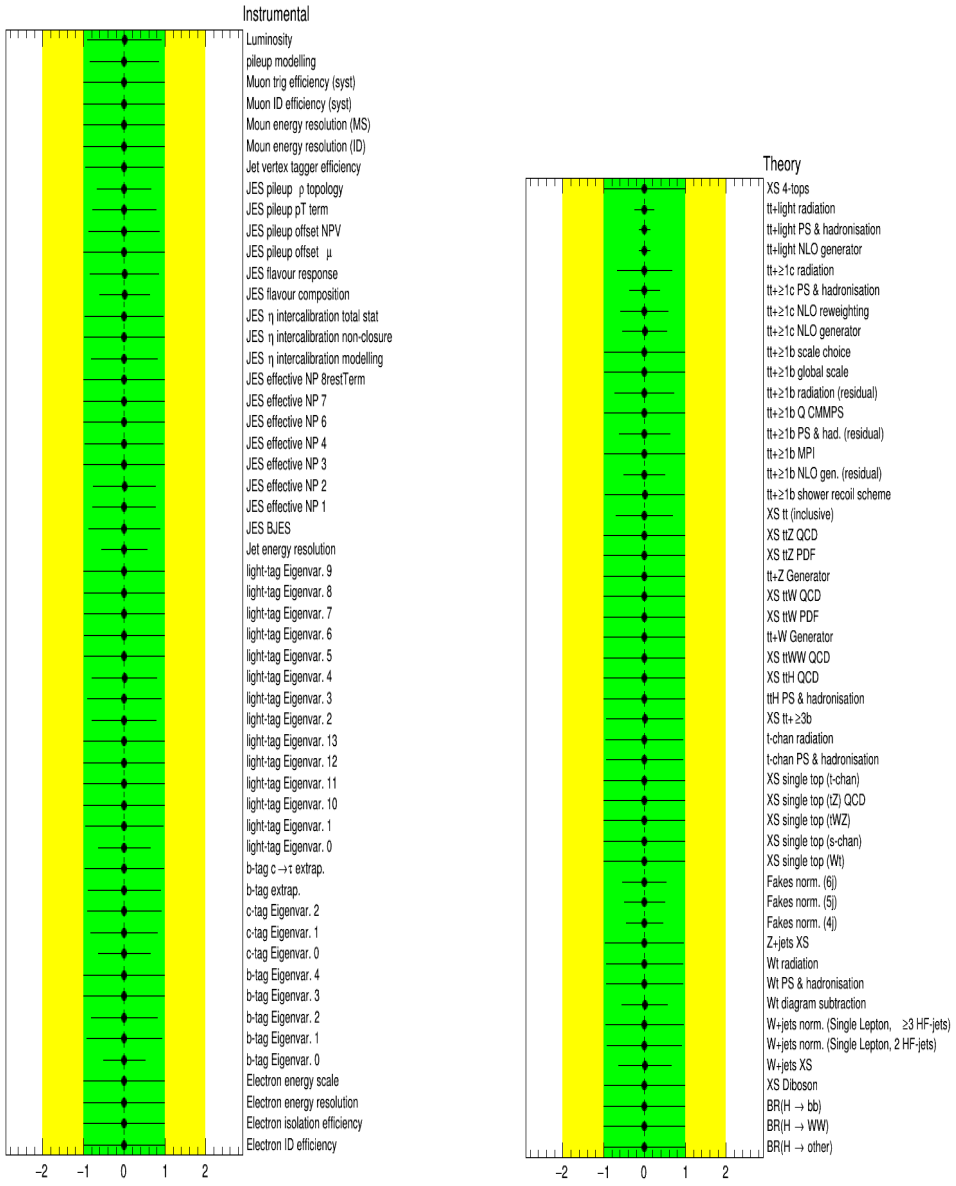


Figure E.1: non-KLFitter nuisance parameters corresponding to the Instrumental (left plot) and Theory (right plot) systematic uncertainties in fits to the Asimov dataset.

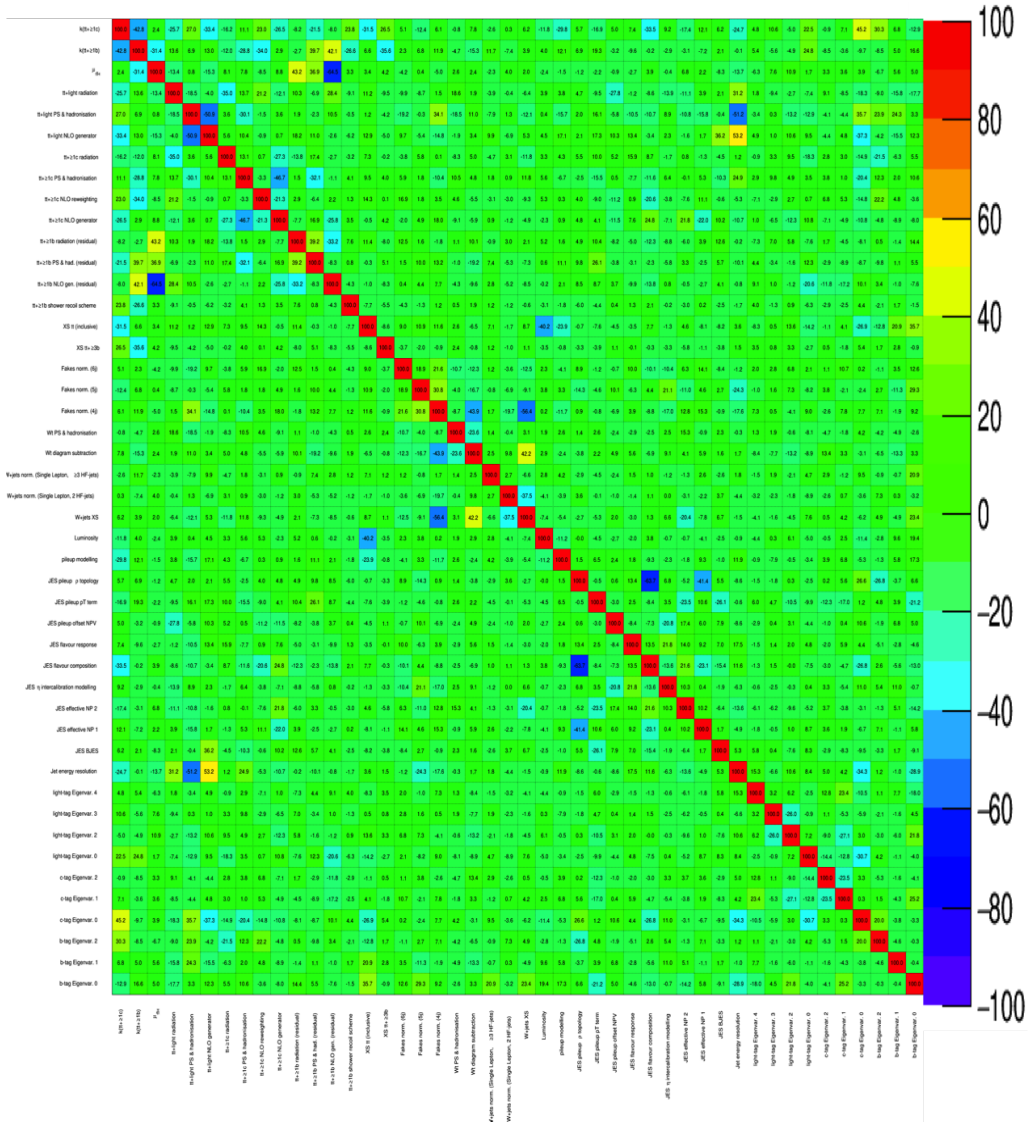


Figure E.2: non-KLFitter correlation matrix between NPs corresponding to the fit to Asimov dataset.

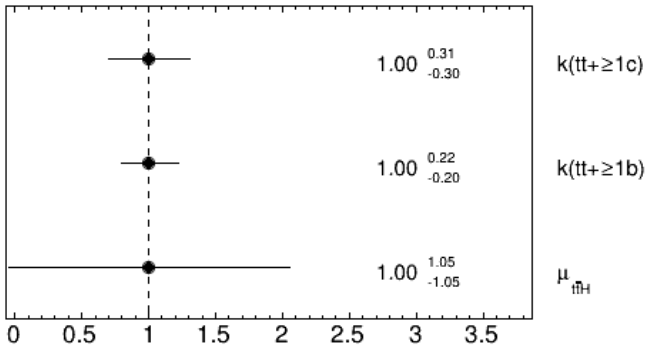


Figure E.3: non-KLFitter normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ and Parameter of interest (signal strength μ) components corresponding to the fit to Asimov dataset.

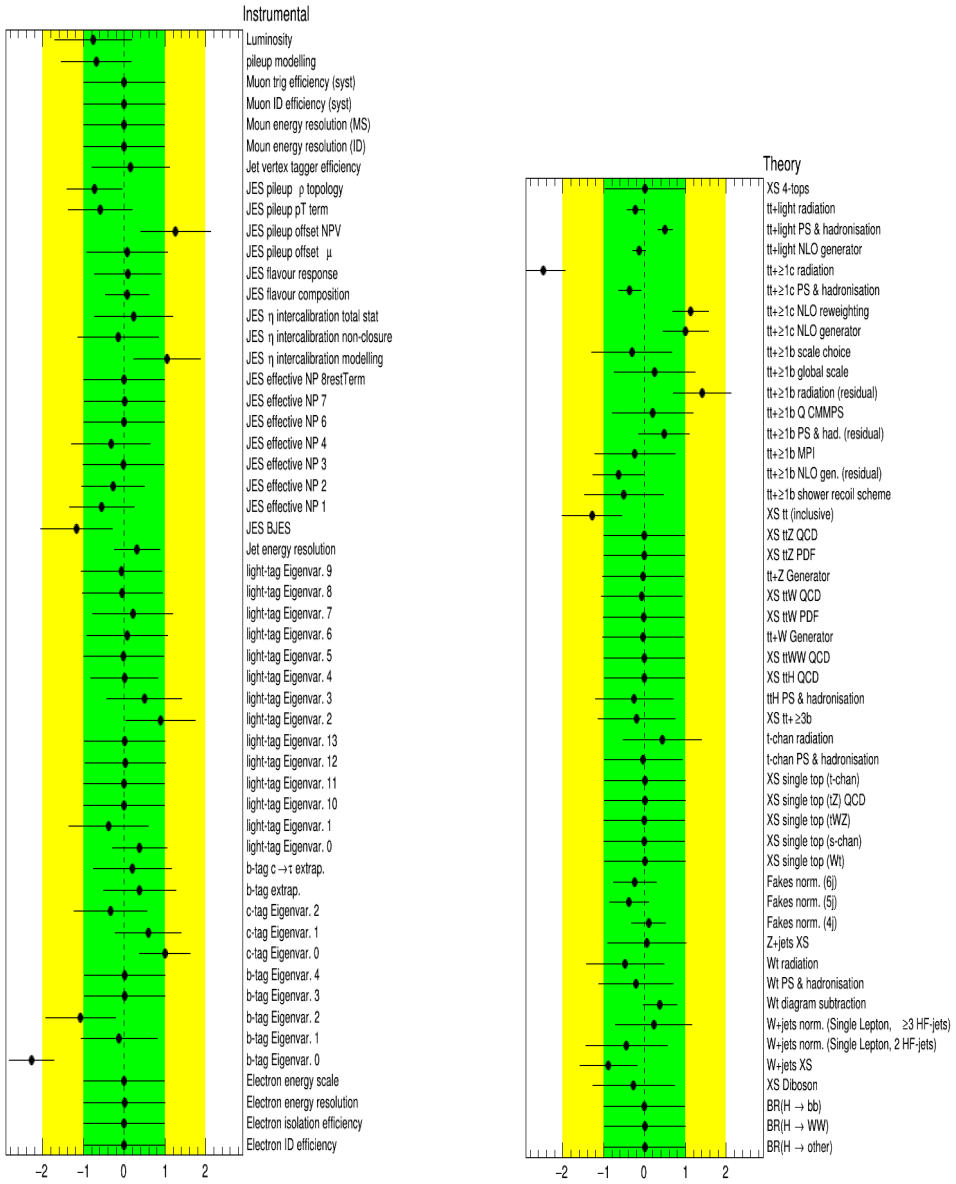


Figure E.4: The non-KLFitter nuisance parameters corresponding to the Instrumental (left plot) and Theory (right plot) systematic uncertainties in fits to data.

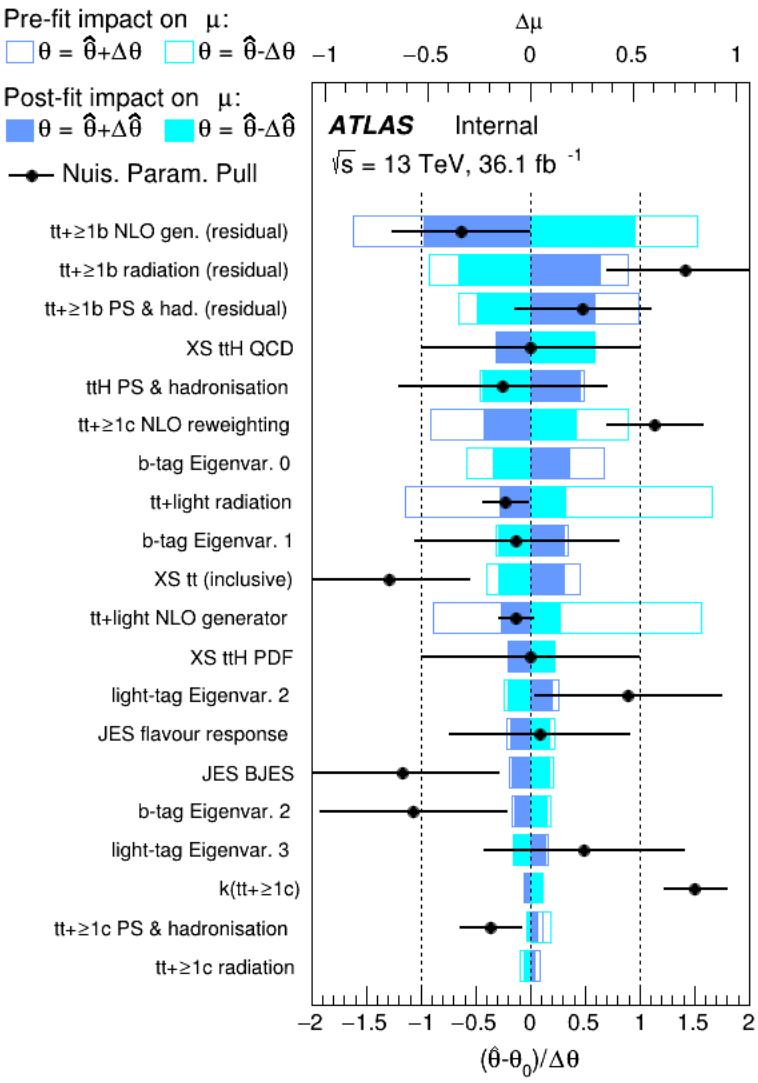


Figure E.5: The fitted values of the non-KLFitter 20 NPs which have the largest effect on the measured signal strength corresponding to the fit to data.

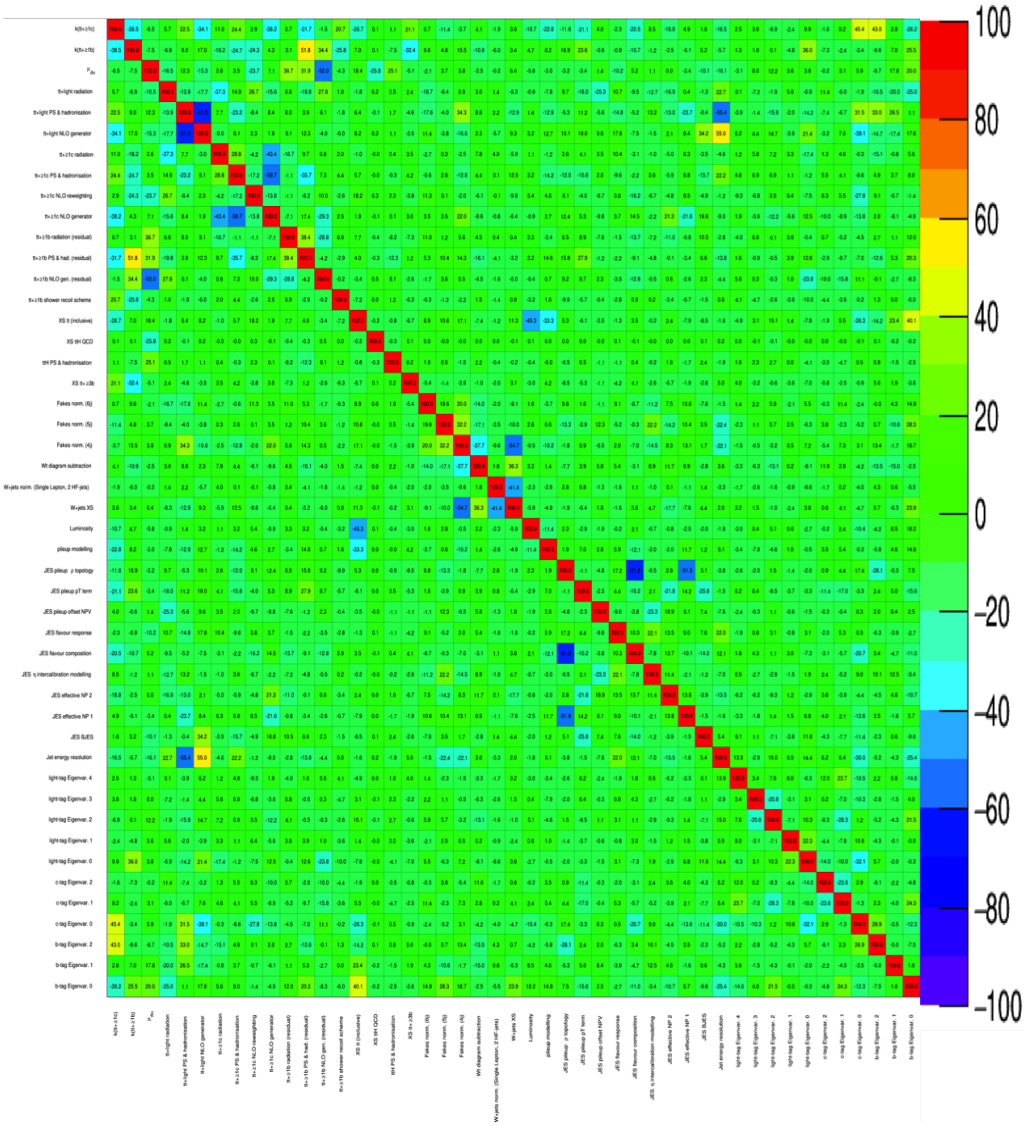


Figure E.6: The non-KLFitter correlation matrix between NPs corresponding to the fit to data.

	= 4j, = 2b	= 4j, = 3b	= 4j, = 4b	= 5j, = 2b	= 5j, = 3b	= 5j, ≥ 4b	≥ 6j, = 2b	≥ 6j, = 3b	≥ 6j, ≥ 4b
Fakes	22 800 ± 11 400	1870 ± 944	76.0 ± 41.8	8730 ± 4370	914 ± 466	0 ± 0	5080 ± 2550	0 ± 0	0 ± 0
$t\bar{t}$ + light	475 000 ± 49 900	43 800 ± 6320	546 ± 293	279 000 ± 41 900	32 000 ± 5900	873 ± 302	176 000 ± 51 200	25 000 ± 7550	1140 ± 617
$t\bar{t}$ + ≥1c	40 900 ± 11 800	8070 ± 2480	275 ± 75.8	36 800 ± 6750	9590 ± 1770	684 ± 240	37 300 ± 9890	12 400 ± 3560	1520 ± 613
$t\bar{t}$ + ≥1b	12 100 ± 2720	5640 ± 1860	443 ± 178	11 500 ± 1780	7890 ± 1810	1270 ± 349	12 600 ± 1920	11 500 ± 2170	3950 ± 1030
$t\bar{t} + V$	107 ± 10.7	29.1 ± 4.07	2.73 ± 0.524	73.1 ± 8.19	19.0 ± 2.40	3.94 ± 0.787	92.1 ± 14.2	44.6 ± 12.1	21.9 ± 8.75
Single top tchan	11 800 ± 1870	882 ± 144	38.5 ± 26.3	3960 ± 782	460 ± 97.9	35.2 ± 11.0	1730 ± 437	309 ± 94.5	48.8 ± 22.6
Single top schan	905 ± 116	67.3 ± 11.1	2.80 ± 0.869	295 ± 47.5	32.4 ± 5.39	1.88 ± 0.677	114 ± 25.6	20.0 ± 4.43	2.71 ± 1.11
Single top Wtchan	19 900 ± 4680	1920 ± 569	58.5 ± 40.3	11 000 ± 3090	1560 ± 510	81.6 ± 35.1	6470 ± 2420	1320 ± 562	162 ± 88.1
Diboson	1120 ± 573	83.0 ± 43.4	4.36 ± 2.55	617 ± 320	64.6 ± 34.5	5.30 ± 3.14	530 ± 282	82.0 ± 43.1	10.0 ± 6.16
$t\bar{t} + W$	303 ± 59.7	34.6 ± 8.44	1.04 ± 0.577	402 ± 69.2	65.4 ± 12.7	3.65 ± 1.24	603 ± 105	145 ± 29.2	18.4 ± 4.98
$t\bar{t} + Z$	332 ± 54.6	64.9 ± 13.5	6.22 ± 1.96	432 ± 57.0	118 ± 18.7	25.0 ± 4.94	761 ± 113	275 ± 42.9	94.7 ± 17.1
W + jets	31 100 ± 14 700	1930 ± 996	34.1 ± 19.1	13 800 ± 6520	1260 ± 624	55.3 ± 29.3	8770 ± 4260	1150 ± 597	125 ± 67.2
Z + jets	5470 ± 2100	367 ± 170	8.36 ± 3.93	2420 ± 964	272 ± 152	11.6 ± 5.23	1570 ± 673	207 ± 87.4	22.2 ± 9.57
$t\bar{t}H$ ($H \to b\bar{b}$)	101 ± 12.8	71.8 ± 11.0	12.9 ± 2.56	133 ± 14.8	138 ± 17.7	48.0 ± 8.86	210 ± 29.3	292 ± 34.8	183 ± 33.7
$t\bar{t}H$ ($H \to WW$)	30.4 ± 5.10	3.66 ± 0.855	0.135 ± 0.144	55.3 ± 6.63	9.05 ± 1.39	0.478 ± 0.269	157 ± 20.2	42.2 ± 6.18	5.72 ± 1.29
$t\bar{t}H$ ($H \to$ other)	31.6 ± 4.63	4.60 ± 0.785	0.273 ± 0.0823	54.3 ± 7.68	10.7 ± 1.71	1.17 ± 0.470	122 ± 17.0	36.9 ± 5.66	6.69 ± 1.30
Total	622 000 ± 62 700	64 800 ± 8840	1510 ± 415	370 000 ± 49 200	54 400 ± 8190	3100 ± 665	252 000 ± 58 800	52 800 ± 11 200	7320 ± 1730
Data 2016	609458	62421	1590	363928	54678	3339	248492	55102	7652

Table E.1: Pre-fit event yields for the non-KLFitter single lepton channel regions fit to data. The three signal regions are highlighted in bold text.

	$= 4$j, $= 2$b	$= 4$j, $= 3$b	$= 4$j, $= 4$b	$= 5$j, $= 2$b	$= 5$j, $= 3$b	$= 5$j, ≥ 4b	≥ 6j, $= 2$b	≥ 6j, $= 3$b	≥ 6j, ≥ 4b
Fakes	24 000 ± 4780	1970 ± 409	79.7 ± 23.6	6910 ± 2180	724 ± 245	0 ± 0	4430 ± 1400	0 ± 0	0 ± 0
$t\bar{t}$ + light	469 000 ± 9650	40 100 ± 1870	491 ± 84.6	275 000 ± 7400	29 900 ± 1760	742 ± 129	175 000 ± 6090	23 600 ± 1720	1010 ± 193
$t\bar{t}$ + $\geq 1c$	46 300 ± 9800	9870 ± 1880	413 ± 70.5	44 900 ± 7920	12 800 ± 2050	1020 ± 165	42 000 ± 6780	16 200 ± 2350	1930 ± 334
$t\bar{t}$ + $\geq 1b$	10 500 ± 2290	5480 ± 1140	411 ± 77.7	9320 ± 2040	7590 ± 1440	1190 ± 196	9850 ± 2130	11 000 ± 1950	3580 ± 439
$t\bar{t} + V$	110 ± 9.45	32.6 ± 3.17	3.49 ± 0.441	73.5 ± 6.40	19.9 ± 1.86	4.67 ± 0.577	87.9 ± 10.9	43.4 ± 11.1	23.8 ± 9.08
Single top tchan	12 200 ± 1380	914 ± 86.8	42.6 ± 24.8	4130 ± 500	521 ± 74.6	43.7 ± 9.34	1820 ± 263	344 ± 75.4	52.9 ± 17.9
Single top schan	945 ± 67.7	72.6 ± 6.83	2.89 ± 0.695	307 ± 23.0	33.6 ± 3.35	2.33 ± 0.576	114 ± 9.85	20.1 ± 2.36	3.16 ± 0.714
Single top Wtchan	20 500 ± 2900	1910 ± 337	44.0 ± 19.5	10 500 ± 1430	1440 ± 248	85.7 ± 24.5	5440 ± 994	1060 ± 249	137 ± 42.4
Diboson	962 ± 528	74.3 ± 41.2	4.11 ± 2.47	485 ± 267	53.9 ± 30.1	4.27 ± 2.68	428 ± 236	67.0 ± 37.1	9.44 ± 5.74
$t\bar{t} + W$	328 ± 58.6	35.3 ± 7.73	0.987 ± 0.524	421 ± 68.1	65.6 ± 11.3	3.50 ± 1.02	603 ± 87.4	140 ± 23.5	17.7 ± 3.86
$t\bar{t} + Z$	356 ± 50.6	71.2 ± 12.4	7.29 ± 1.90	446 ± 54.6	123 ± 17.0	28.0 ± 4.23	744 ± 89.4	267 ± 33.1	103 ± 13.7
W + jets	17 900 ± 6150	1180 ± 536	21.5 ± 10.9	7970 ± 2780	743 ± 294	43.4 ± 19.1	4930 ± 1720	746 ± 319	84.0 ± 36.5
Z + jets	5540 ± 1890	356 ± 141	9.58 ± 3.78	2410 ± 826	234 ± 125	13.5 ± 5.20	1510 ± 533	203 ± 71.1	23.4 ± 8.57
$t\bar{t}H$ ($H \to b\bar{b}$)	315 ± 91.1	263 ± 72.8	50.9 ± 13.6	374 ± 112	475 ± 133	182 ± 49.0	507 ± 159	886 ± 257	660 ± 179
$t\bar{t}H$ ($H \to WW$)	103 ± 29.9	12.3 ± 3.39	0.521 ± 0.450	186 ± 53.7	28.9 ± 8.97	1.27 ± 0.848	493 ± 141	129 ± 37.0	17.5 ± 4.97
$t\bar{t}H$ ($H \to$ other)	109 ± 31.2	15.1 ± 4.28	0.819 ± 0.269	185 ± 52.2	34.6 ± 10.7	3.56 ± 1.31	381 ± 110	112 ± 32.3	21.1 ± 6.05
Total	609 000 ± 1270	62 400 ± 461	1580 ± 59.8	364 000 ± 976	54 800 ± 342	3360 ± 79.1	248 000 ± 719	54 700 ± 349	7680 ± 124
Data 2016	609458	62421	1590	363928	54678	3339	248492	55102	7652

Table E.2: Post-fit event yields for the non-KLFitter single lepton channel regions fit to data. The three signal regions are highlighted in bold text.

Bibliography

[1] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716 (CERN-PH-EP-2012-218):1–29. 29 p, Aug 2012. arXiv: 1207.7214 [hep-ph].

[2] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the cms experiment at the lhc. *Physics Letters B*, 716(1):30 – 61, 2012. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2012.08.021. URL `http://www.sciencedirect.com/science/article/pii/S0370269312008581`.

[3] G Arnison et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$= 540 GeV. *Physics Letters B*, 122(1):103–116, 1983.

[4] G Arnison et al. Experimental observation of lepton pairs of invariant mass around 95 GeV at the CERN SPS collider. *Physics Letters B*, 126(5):398–410, 1983.

[5] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi: 10.1103/PhysRevLett.13.321. URL `https://link.aps.org/doi/10.1103/PhysRevLett.13.321`.

[6] D. De Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. Technical Report FERMILAB-FN-1025-T, Oct 2016. URL `https://cds.cern.ch/record/2227475`. web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.

[7] ATLAS Collaboration et al. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$= 8 TeV with the ATLAS detector. 2015. arXiv: 1503.05066 [hep-ph].

[8] Georges Aad et al. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$

TeV with the ATLAS detector. Technical Report ATL-COM-PHYS-2017-079, CERN, Geneva, Feb 2017. URL https://cds.cern.ch/record/2244360.

[9] S Dittmaier et al. Handbook of LHC Higgs cross sections: 1. inclusive observables. 2011. arXiv: 1101.0593 [hep-ph].

[10] S Heinemeyer et al. Handbook of LHC Higgs cross sections: 3. Higgs properties. 2013. arXiv: 1307.1347 [hep-ph].

[11] Andreas Hoecker et al. TMVA-Toolkit for multivariate data analysis. 2007. arXiv: physics/0703039.

[12] Abdus Salam. Gauge Unification of Fundamental Forces. *Rev. Mod. Phys.*, 52:525–538, 1980. doi: 10.1103/RevModPhys.52.525. [,306(1980)].

[13] Francis Halzen and Alan D Martin. *Quark & Leptons: An Introductory Course In Modern Particle Physics*. John Wiley & Sons, 2008.

[14] W Noel Cottingham and Derek A Greenwood. *An introduction to the standard model of particle physics*. Cambridge university press, 2007.

[15] John F Donoghue et al. *Dynamics of the standard model*, volume 35. Cambridge university press, 2014.

[16] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi: 10.1088/1674-1137/38/9/090001.

[17] I. J. R. Aitchison and A. J. G. Hey. *Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory*. CRC Press, Bristol, UK, 2012. ISBN 9781466513075. URL http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.3.F5A34::2012:V2.

[18] C. Amsler et al. Review of Particle Physics. *Physics Letters B*, 667:414, 2008.

[19] I. Van Vulpen. The Standard Model Higgs Boson. *Part of the Lecture Particle Physics II, UvA Particle Physics Master 2011-2012*, available at http://master.particles.nl/LectureNotes/2011-PPII-Higgs.pdf:22, 2012.

[20] Chien-Shiung Wu et al. Experimental test of parity conservation in beta decay. *Physical review*, 105(4):1413, 1957.

[21] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. doi: 10.1103/PhysRevLett.19.1264. URL `https://link.aps.org/doi/10.1103/PhysRevLett.19.1264`.

[22] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.

[23] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012.

[24] ATLAS collaboration et al. Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC. *Physics Letters B*, 726(1):88–119, 2013.

[25] Atlas Collaboration et al. Evidence for the spin-0 nature of the Higgs boson using Atlas data. *Physics Letters B*, 726(1):120–144, 2013.

[26] Georges Aad et al. Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector. Technical report, ATLAS-HIGG-2013-32-003, 2015.

[27] CMS Collaboration et al. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. 2014. arXiv: 1412.8662 [hep-ph].

[28] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s}$= 7 and 8 TeV with the ATLAS and CMS Experiments. *Physical review letters*, 114(19):191803, 2015.

[29] F Maltoni et al. Measuring the top-quark Yukawa coupling at hadron colliders via $t\bar{t}H$, H$\rightarrow w^+w^-$. *Physical Review D*, 66(3):034022, 2002.

[30] V Drollinger, Th Muller, and D Denegri. Searching for Higgs Bosons in Association with Top Quark Pairs in the $H \rightarrow b\bar{b}$ Decay Mode. 2001. arXiv: 0111312 [hep-ph].

[31] Victor Ilisie. SM Higgs Decay and Production Channels. *Universidad de Valencia, Septiembre*, 2011.

[32] V Ilisie. SM Higgs Decay and Production Channels. *Universidad de Valencia, Septiembre*, 2011.

[33] S Heinemeyer et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group. Technical Report CERN-2013-004. CERN-2013-004, Jul 2013. URL https://cds.cern.ch/record/1559921. Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[34] Claudia Bertella. *Probing top quark and Higgs boson production in multi-jet events at the LHC with the ATLAS detector*. PhD thesis, Aix-Marseille Université, 2013.

[35] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08001.

[36] D Brandt et al. Accelerator physics at LEP. *Reports on Progress in Physics*, 63(6):939, 2000.

[37] The ATLAS Collaboration et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[38] The CMS Collaboration et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[39] The ALICE Collaboration et al. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[40] The LHCb Collaboration et al. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, 2008. URL http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[41] S Feher and J Strait. Estimated inner triplet quadrupole length and aperture for really large hadron colliders of E (beam)= 30-TeV, 60-TeV, 100-TeV. *eConf*, 960625(SNOWMASS-1996-ACC042):ACC042, 1996.

[42] Werner Herr and B Muratori. Concept of luminosity. 2006. URL http://cds.cern.ch/record/941318.

[43] Federico Roncarolo. *Accuracy of the transverse emittance measurements of the CERN large hadron collider*. PhD thesis, Milan Polytechnic, 2006.

[44] HiLumi LHC Collaboration et al. HL-LHC Preliminary Design Report. *FP7 High Luminosity Large Hadron Collider Design Study, CERN-ACC-2014*, 300, 2014.

[45] Kinematics, Cross-Section Formulae, and Plots. *Physics Letters B*, 667(1): 340 – 370, 2008. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2008.07.031. URL http://www.sciencedirect.com/science/article/pii/S0370269308008563. Review of Particle Physics.

[46] Tao Han. Collider phenomenology: Basic knowledge and techniques. In *Physics in D $\geq$ 4. Proceedings, Theoretical Advanced Study Institute in elementary particle physics, TASI 2004, Boulder, USA, June 6-July 2, 2004*, pages 407–454, 2005. doi: 10.1142/9789812773579_0008. arXiv: hep-ph/0508097.

[47] ATLAS inner detector: Technical design report. Vol. 1. 1997.

[48] Georges Aad et al. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(8):S08003–S08003, 2008.

[49] The ATLAS collaboration. Operation and performance of the ATLAS semiconductor tracker. *Journal of Instrumentation*, 9(08):P08009, 2014. URL http://stacks.iop.org/1748-0221/9/i=08/a=P08009.

[50] Anatoli Romaniouk et al. Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties. Technical Report ATL-COM-INDET-2015-041, CERN, Geneva, Jun 2015. URL https://cds.cern.ch/record/2021497.

[51] Laura Barranco Navarro. Alignment of the ATLAS Inner Detector in the LHC RunII. Technical Report ATL-PHYS-PROC-2015-190, CERN, Geneva, Dec 2015. URL https://cds.cern.ch/record/2114708.

[52] M Capeans et al. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013. ATLAS-TDR-19, Sep 2010. URL https://cds.cern.ch/record/1291633.

[53] Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s}$ = 13 TeV. Technical Report ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037683.

[54] E Abat et al. The Atlas TRT barrel detector. *Journal of Instrumentation*, 3 (02):P02014, 2008.

[55] E Abat et al. The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance. *Journal of Instrumentation*, 3 (02):P02013, 2008.

[56] L. Fayard. Transition radiation. *Instrumentation en physique nucleaire et en physique des particules*, available at http://hal.in2p3.fr/cel-00645585/document(cel-00645585):22, 26-30 septembre 1988.

[57] W. M. Yao et al. Review of Particle Physics. *J. Phys.*, G33:1–1232, 2006. doi: 10.1088/0954-3899/33/1/001.

[58] Michele Livan and Richard Wigmans. Misconceptions about calorimetry. *Instruments*, 1(1):3, 2017.

[59] Georges Aad et al. Performance of the ATLAS Trigger System in 2010. *Eur. Phys. J.*, C72:1849, 2012. doi: 10.1140/epjc/s10052-011-1849-1. arXiv: 1110.1530 [hep-ph].

[60] Allen Caldwell et al. BAT - The Bayesian analysis toolkit. *Computer Physics Communications*, 180(11):2197–2209, 2009.

[61] Georges Aad et al. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATL-COM-PHYS-2016-116, CERN, Geneva, Feb 2016. URL `https://cds.cern.ch/record/2130799`.

[62] V Barger et al. Event shape criteria for single-lepton top-quark signals. *Physical Review D*, 48(9):R3953, 1993.

[63] David Ciupke. Study of BDT training configurations with an application to the Z/H τ τ *ee* Analysis, 2012.

[64] Glen Cowan et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, Feb 2011. ISSN 1434-6052. doi: 10.1140/epjc/s10052-011-1554-0. URL `https://doi.org/10.1140/epjc/s10052-011-1554-0`.

[65] T Cornelissen et al. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar 2007. URL `https://cds.cern.ch/record/1020106`.

[66] Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment. Technical Report ATLAS-CONF-2010-069, CERN, Geneva, Jul 2010. URL https://cds.cern.ch/record/1281344.

[67] Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment. Technical Report ATLAS-CONF-2012-042, CERN, Geneva, Mar 2012. URL https://cds.cern.ch/record/1435196.

[68] The Optimization of ATLAS Track Reconstruction in Dense Environments. Technical Report ATL-PHYS-PUB-2015-006, CERN, Geneva, Mar 2015. URL https://cds.cern.ch/record/2002609.

[69] b-tagging in dense environments. Technical Report ATL-PHYS-PUB-2014-014, CERN, Geneva, Aug 2014. URL https://cds.cern.ch/record/1750682.

[70] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical Report ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016. URL https://cds.cern.ch/record/2157687.

[71] Lepton isolation recommendations. URL https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/IsolationSelectionTool.

[72] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$= 13 TeV. *The European Physical Journal C*, 76(5):1–30, 2016.

[73] Matteo Cacciari et al. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[74] Georges Aad et al. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}$= 7 TeV. *The European Physical Journal C*, 73 (3):2304, 2013.

[75] Georges Aad et al. Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s}$= 7 TeV with the ATLAS detector. *The European Physical Journal C*, 75(1):17, 2015.

[76] ATLAS Collaboration et al. Jet reconstruction and performance using particle flow with the ATLAS Detector. 2017. arXiv: 1703.10485 [hep-ph].

[77] Aliaksei Hrynevich. ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2. Technical Report ATL-PHYS-PROC-2017-045. 06, CERN, Geneva, May 2017. URL https://cds.cern.ch/record/2263777.

[78] W Lampl et al. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008. URL https://cds.cern.ch/record/1099735.

[79] Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037613.

[80] Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector. Technical Report ATLAS-CONF-2015-029, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037702.

[81] Bad Loose selection. URL https://twiki.cern.ch/twiki/bin/view/Sandbox/EgammaPhysCalib.

[82] ATLAS Collaboration et al. Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector. Number ATLAS-CONF-2015-029, 2015. URL https://cds.cern.ch/record/2037702.

[83] Georges Aad et al. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$= 8 TeV using the ATLAS detector. *The European Physical Journal C*, 76(11):581, 2016.

[84] Expected performance of the ATLAS b-tagging algorithms in Run-2. Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2015. URL http://cds.cern.ch/record/2037697.

[85] Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions at 8 TeV. Technical Report ATLAS-CONF-2013-082, CERN, Geneva, Aug 2013. URL https://cds.cern.ch/record/1570993.

[86] BMM Allbrooke et al. Performance of algorithms that reconstruct missing transverse momentum in center of mass 8 TeV proton–proton collisions in the ATLAS detector. *European Physical Journal C: Particles and Fields*, 77(6): a241, 2017.

[87] Data derivations (25ns). URL availableathttps://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopDerivationMC15cList.

[88] Trigger Menu in 2016. Technical Report ATL-DAQ-PUB-2017-001, CERN, Geneva, Jan 2017. URL https://cds.cern.ch/record/2242069.

[89] ATLAS Collaboration. Performance of the ATLAS trigger system in 2015. *Eur. Phys. J. C*, 77:317, 2017.

[90] Jeannine Wagner-Kuhr. Experimental Studies of Top Quark Production. 2016. arXiv: 1606.02936 [hep-ph].

[91] CB Jackson et al. Higgs in space! *Journal of Cosmology and Astroparticle Physics*, 2010(04):004, 2010.

[92] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. doi: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[93] ATLAS RunI Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014. URL https://cds.cern.ch/record/1966419.

[94] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, 2004.

[95] Stefano Frixione et al. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007 (11):070, 2007.

[96] Simone Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6):43, 2010.

[97] Torbjörn Sjöstrand et al. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11):852–867, 2008.

[98] Fabio Cascioli et al. NLO matching for production with massive b-quarks. *Physics Letters B*, 734:210–214, 2014.

[99] Richard D Ball et al. Parton distributions for the LHC RunII. *Journal of High Energy Physics*, 2015(4):40, 2015.

[100] David J Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152–155, 2001.

[101] Michał Czakon and Alexander Mitov. Top++: a program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185(11):2930–2938, 2014.

[102] Matteo Cacciari et al. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Physics Letters B*, 710 (4):612–622, 2012.

[103] Peter Bärnreuther et al. Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t}+$ X. *Phys. Rev. Lett.*, 109 (CERN-PH-TH-2012-092):132001, 2012.

[104] Michal Czakon and Alexander Mitov. NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *Journal of High Energy Physics*, 2012(12):54, 2012.

[105] Michal Czakon and Alexander Mitov. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. *Journal of High Energy Physics*, 2013(1):80, 2013.

[106] Michal Czakon et al. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, 2013. doi: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph].

[107] Tanju Gleisberg et al. Event generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007, 2009.

[108] Steffen Schumann and Frank Krauss. A parton shower algorithm based on Catani-Seymour dipole factorisation. *Journal of High Energy Physics*, 2008 (03):038, 2008.

[109] Stefan Höche et al. QCD matrix elements and truncated showers. *Journal of High Energy Physics*, 2009(05):053, 2009.

[110] Fabio Cascioli et al. Scattering amplitudes with open loops. *Physical review letters*, 108(11):111601, 2012.

[111] A Bredenstein et al. NLO QCD corrections to top anti-top bottom anti-bottom production at the LHC. In *Proceedings of the 35th International Conference*

of High Energy Physics (ICHEP 2010). July 22-28, 2010. Paris, France. Published online at http://pos. sissa. it/cgi-bin/reader/conf. cgconfid= 120", id. 105, 2010.

[112] G Bevilacqua et al. Assault on the NLO wishlist: pp→ $t\bar{t}b\bar{b}$. *Journal of High Energy Physics*, 2009(09):109, 2009.

[113] Studies of $t\bar{t}+c\bar{c}$ production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment. Technical Report ATL-PHYS-PUB-2016-011, CERN, Geneva, May 2016. URL `https://cds.cern.ch/record/2153876`.

[114] Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *Journal of High Energy Physics*, 2008(12):039, 2008.

[115] Stefan Hoeche et al. QCD matrix elements+ parton showers. the NLO case. *Journal of High Energy Physics*, 2013(4):27, 2013.

[116] J Butterworth et al. Single Boson and Diboson Production Cross Sections in *pp* Collisions at $\sqrt{s}$=7 TeV. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Aug 2010. URL `https://cds.cern.ch/record/1287902`.

[117] Frederic Derue. Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s}$=13 TeV with AnalysisTop-2.3.41. Technical Report ATL-COM-PHYS-2016-198, CERN, Geneva, Feb 2016. URL `https://cds.cern.ch/record/2135116`.

[118] Pierre Artoisenet et al. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *Journal of High Energy Physics*, 2013(3):15, Mar 2013. ISSN 1029-8479. doi: 10.1007/JHEP03(2013)015. URL `https://doi.org/10.1007/JHEP03(2013)015`.

[119] Peter Z. Skands. Tuning Monte Carlo generators: The Perugia tunes. *Phys. Rev. D*, 82:074018, Oct 2010. doi: 10.1103/PhysRevD.82.074018. URL `https://link.aps.org/doi/10.1103/PhysRevD.82.074018`.

[120] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Physical Review D*, 82(5):054018, 2010.

[121] Nikolaos Kidonakis. Next-to-next-to-leading logarithm resummation for s-channel single top quark production. *Physical Review D*, 81(5):054028, 2010.

[122] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Physical Review D*, 83 (9):091503, 2011.

[123] Stefano Frixione et al. Single-top hadroproduction in association with a W boson. *Journal of High Energy Physics*, 2008(07):029, 2008.

[124] Richard D Ball et al. Parton distributions with LHC data. *Nuclear Physics B*, 867(2):244–289, 2013.

[125] W Beenakker et al. Higgs radiation off top quarks at the Tevatron and the LHC. *Physical Review Letters*, 87(20):201805, 2001.

[126] S Dawson et al. Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections. 2003. arXiv: 0305087 [hep-ph].

[127] D. De Florian et al. Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector. Technical report, 2016.

[128] L. Reina and S. Dawson. Next-to-Leading Order Results for $t\bar{t}H$ Production at the Tevatron. *Phys. Rev. Lett.*, 87:201804, Oct 2001. doi: 10.1103/PhysRevLett.87.201804. URL `https://link.aps.org/doi/10.1103/PhysRevLett.87.201804`.

[129] W Beenakker et al. NLO QCD corrections to $t\bar{t}H$ production in hadron collisions. *Nuclear Physics B*, 653(1):151–203, 2003.

[130] Abdel Djouadi et al. HDECAY: A program for higgs boson decays in the standard model and its supersymmetric extension. *Computer Physics Communications*, 108(1):56–74, 1998.

[131] A Bredenstein et al. Precise predictions for the Higgs-boson decay $H \to$ W W/Z Z 4 leptons. *Physical Review D*, 74(1):013004, 2006.

[132] Stefano Actis et al. NNLO computational techniques: the cases H$\to \gamma\gamma$ and $h \to gg$. *Nuclear Physics B*, 811(1):182–273, 2009.

[133] A Denner et al. Standard model Higgs-boson branching ratios with uncertainties. *The European Physical Journal C*, 71(9):1753, 2011.

[134] Johannes Erdmann et al. A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 748:18–25, 2014.

[135] Torbjörn Sjöstrand et al. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006.

[136] Catherine Bernaciak et al. Fox-Wolfram moments in Higgs physics. *Phys. Rev. D*, 87:073014, Apr 2013. doi: 10.1103/PhysRevD.87.073014. URL https://link.aps.org/doi/10.1103/PhysRevD.87.073014.

[137] V. Barger et al. Event shape criteria for single-lepton top-quark signals. *Phys. Rev. D*, 48:R3953–R3956, Nov 1993. doi: 10.1103/PhysRevD.48.R3953. URL https://link.aps.org/doi/10.1103/PhysRevD.48.R3953.

[138] Jet energy measurement and systematic uncertainties using tracks for jets and for b-quark jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS detector. Technical Report ATLAS-CONF-2013-002, CERN, Geneva, Jan 2013. URL https://cds.cern.ch/record/1504739.

[139] Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2015-057, CERN, Geneva, Oct 2015. URL https://cds.cern.ch/record/2059846.

[140] Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8TeV$. Technical Report ATLAS-CONF-2015-017, CERN, Geneva, Apr 2015. URL https://cds.cern.ch/record/2008678.

[141] Micha Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185(11):2930 – 2938, 2014. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2014.06.021. URL http://www.sciencedirect.com/science/article/pii/S0010465514002264.

[142] Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13$ Te V. Technical Report ATL-PHYS-PUB-2016-004, CERN, Geneva, Jan 2016. URL https://cds.cern.ch/record/2120417.

[143] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. *Phys. Rev. D*, 83:091503, May 2011. doi: 10.1103/PhysRevD.83.091503. URL https://link.aps.org/doi/10.1103/PhysRevD.83.091503.

[144] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. *Phys. Rev. D*, 82:054018, Sep 2010.

doi: 10.1103/PhysRevD.82.054018. URL https://link.aps.org/doi/10.1103/PhysRevD.82.054018.

[145] Nikolaos Kidonakis. Next-to-next-to-leading logarithm resummation for s-channel single top quark production. *Phys. Rev. D*, 81:054028, Mar 2010. doi: 10.1103/PhysRevD.81.054028. URL https://link.aps.org/doi/10.1103/PhysRevD.81.054028.

[146] S.Frixione et al. Single-top hadroproduction in association with a W boson. *Journal of High Energy Physics*, 2008(07):029, 2008. URL http://stacks.iop.org/1126-6708/2008/i=07/a=029.

[147] John M. Campbell and R. Keith Ellis. $t\bar{t}W^{\pm}$ production and decay at NLO. *Journal of High Energy Physics*, 2012(7):52, Jul 2012. ISSN 1029-8479. doi: 10.1007/JHEP07(2012)052. URL https://doi.org/10.1007/JHEP07(2012)052.

[148] S Dawson et al. Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections. *Phys. Rev. D*, 68:034022, Aug 2003. doi: 10.1103/PhysRevD.68.034022. URL https://link.aps.org/doi/10.1103/PhysRevD.68.034022.

[149] W. Beenakker et al. NLO QCD corrections to $t\bar{t}H$ production in hadron collisions. *Nuclear Physics B*, 653(1):151 – 203, 2003. ISSN 0550-3213. doi: https://doi.org/10.1016/S0550-3213(03)00044-0. URL http://www.sciencedirect.com/science/article/pii/S0550321303000440.

[150] Walter W et al. Higgs-boson production at large transverse momentum in quantum chromodynamics. *Phys. Rev. D*, 19:941–944, Feb 1979. doi: 10.1103/PhysRevD.19.941. URL https://link.aps.org/doi/10.1103/PhysRevD.19.941.

[151] Yu Zhang et al. QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider. *Physics Letters B*, 738:1 – 5, 2014. ISSN 0370-2693. doi: https://doi.org/10.1016/j.physletb.2014.09.022. URL http://www.sciencedirect.com/science/article/pii/S0370269314006728.

[152] S Frixione et al. Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons. *Journal of High Energy Physics*, 2015(6):184, Jun 2015. ISSN 1029-8479. doi: 10.1007/JHEP06(2015)184. URL https://doi.org/10.1007/JHEP06(2015)184.

[153] Glen Cowan et al. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, 2011.

[154] Morad Aaboud et al. Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Submitted to: Phys. Rev. D*, 2017. arXiv: 1712.08895 [hep-ex].

[155] P Watkins. Proceedings of the School for Young High Energy Physicists. Rutherford Appleton Laboratory. Technical report, Council for the Central Lab. of the Research Councils (CLRC), 2001.

[156] Georges Aad et al. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *The European Physical Journal C*, 74(11):3130, 2014.